

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2003, 06:43:25 ; Search time 3256 Seconds

(without alignments)
10627.529 Million cell updates/sec

Title: US-09-857-613A-27

Perfect score: 1189

Sequence: 1 ggcacatggccacccgtgtgtga.....aaaaaaaaaaaaaaaaaaaaa 1189

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenBank:*

- 1: gb_ba:*
- 2: gb_hlg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pt:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_ot:*
- 22: em_ov:*
- 23: em_ph:*
- 24: em_pl:*
- 25: em_ro:*
- 26: em_sts:*
- 27: em_un:*
- 28: em_vl:*
- 29: em_ba:*
- 30: em_hlg:*
- 31: em_in:*
- 32: em_mu:*
- 33: em_om:*
- 34: em_ot:*
- 35: em_ov:*
- 36: em_ph:*
- 37: em_pl:*
- 38: em_ro:*
- 39: em_sts:*
- 40: em_un:*
- 41: em_vl:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	456.8	38.4	1369	AF213481	AF213481 Perilla frutescens gamma-tocopherol methyltransferase (TMT) mRNA, complete cds.
2	445.2	37.4	1303	AY049258	AY049258 Arabidops
3	443.6	37.3	1047	AY090280	AY090280 Arabidops
4	442	37.2	1350	AF104220	AF104220 Arabidops
5	440.4	37.0	1296	AY087138	AY087138 Arabidops
6	440.4	16.2	118335	AC006193	AC006193 Arabidops
7	192.2	12.6	343550	AP003587	AP003587 Nostoc sp
8	149.4	12.4	134199	SYCSIRF	SYCSIRF AP003994 Oryza sat
9	147	9.3	125422	AF003994	AF003994 Sorghum b
10	110	8.4	108553	AF527809	AF527809 Sorghum b
11	99.4	8.4	930	AX073663	AX073663 Sequence
12	76.6	6.4	957	AX073657	AX073657 Sequence
13	76.6	6.4	957	AX417730	AX417730 Sequence
14	76.6	6.4	974	AX073659	AX073659 Sequence
15	76.6	6.4	145709	D90914	D90914 Synechocyst
16	76.6	5.8	334520	AP003588	AP003588 Nostoc sp
17	68.8	5.7	298750	AP005375	AP005375 Thermosyn
18	67.4	4.3	218856	AC099282	AC099282 Rattus no
19	51.6	4.3	7218	166494	166494 Sequence 14
20	51	4.3	293545	AC098276	AC098276 Rattus no
21	50.6	4.2	10969	SPAC19E9	SPAC19E9 S. pombe chr
22	50	4.1	1665	AB025584	AB025584 Dictyoste
23	49	4.1	234488	AC109698	AC109698 Rattus no
24	49	4.1	1851	AB064996	AB064996 Macaca fa
25	48.4	4.1	25681	SAE41559	SAE41559 Saccharot
26	48.2	4.1	32070	AC116983	AC116983 Dictyoste
27	48.2	4.0	67970	PFMALIP3	PFMALIP3 Al031746 Plasmodi
28	48	4.0	180260	AC107168	AC107168 Rattus no
29	47.8	4.0	1410	E07846	E07846 DNA sequenc
30	47.8	4.0	5998	AB032524	AB032524 Streptomy
31	47.8	4.0	7143	AX345885	AX345885 Sequence
32	47.6	4.0	161557	AC009826	AC009826 Homo sapi
33	47.6	4.0	193024	AC073855	AC073855 Homo sapi
34	47.6	4.0	193024	AC073855	AC073855 Homo sapi
35	47.6	4.0	107289	AC116923	AC116923 Dictyoste
36	47.2	4.0	12381	AX006889	AX006889 Sequence
37	46.8	3.9	128124	AC113915	AC113915 Rattus no
38	46.8	3.9	349980	AX344572	AX344572 Gentiana
39	46.8	3.9	2180	AB028665	AB028665 Rattus no
40	46.6	3.9	239685	AC096310	AC096310 Caenorhab
41	46.4	3.9	298406	CEY7588A	CEY7588A Dictyoste
42	46.4	3.9	105470	AC116306	AC116306 Homo sapi
43	46.2	3.9	154611	AC068893	AC068893 Rattus no
44	46.2	3.9	170742	AC109078	AC109078 Rattus no
45	46.2	3.9	170742	AC109078	AC109078 Rattus no

ALIGNMENTS

RESULT 1
AF213481 1369 bp mRNA linear PLN 02-DEC-200
LOCUS Perilla frutescens gamma-tocopherol methyltransferase (TMT) mRNA,
DEFINITION complete cds.
ACCESSION AF213481
VERSION AF213481
KEYWORDS
SOURCE
ORGANISM Perilla frutescens
Perilla frutescens
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; eusterids I; Lamiales; Lamiales; Perilla.
REFERENCE
1 (bases 1 to 1369)
Kim, K.-H., Hwang, S.-K. and Hwang, Y.-S.

TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	source
Cloning of <i>Perilla</i> gamma-tocopherol methyltransferase	Unpublished	2 (bases 1 to 1369)	Kim, K.-H., Hwang, S.-K. and Hwang, Y.-S.	Submitted (09-DEC-1999)	Division of Biochemistry, National Institute of Agricultural Science and Technology, 249 Seodun-dong, Kwonson-gu, Suwon, Kyunggi-do 441-707, Republic of Korea	Location/Qualifiers	1. .1369
							/organism="Perilla frutescens"
							/cultivar="Okdong"
							/db_xref="taxon:46386"
							/rissue_type="developing seed"
							1. .1369
							/gene="TMT"
							89. .1198
							/gene="TMT"
							/codon_start=1
							/product="gamma-tocopherol methyltransferase"
							/protein_id="AAL3693.1"
							/db_xref="GI:17224293"
							/translation="MAEAVPFGICTGMRBGVAPPTYNISIKPATALLVCCITKTS
							ITSPESDLRTGRARRRPTNMAAABEMEREMETLEFVGLAEFYDESSGVENITWGBH
							MHGFYEPADVLSIDHRAADIRMI EESLRFAFSPIYTTETKPNIVDVGIGSSRS
							YLARKYGIKLSRAITLSSPVQAOARQADADQNGKVSFEVADALNPFEQGFDFIV
							WSMEGHEMPDKKFFVNELVPAAGGRIILVTYCHRDLSPESSILROEEDLINKITC
							SAYILPAMCSTADYVKILKLDLSMEDIKSMDSHVAPFPWPAIVISALTWKQITSLNS
							GKMTIRGAMVNPIMIEGKKKVIKFIATITCKRPAS"
BASE COUNT	380 a	306 c	332 g	351 t			
ORIGIN							
Query Match	38.4%	Score 456.8	DB 8	Length 1369			
Best Local Similarity	71.2%	Pred. No. 3.4e-105					
Matches 634:	Conservative	0;	Mismatches 247;	Indels 9;	Gaps 2;		
115	CAAGAAGATGACAAAGAACAGCTGCAGAGGGAATCGCAGAGTTTACGACGAGTCGTC	234					
Db	304 GATGGAGACGGAGATGAGGACCTTCGTAAGAGGATGCGGAGTTCTPAGAGAGTCTTC	363					
OY	235 TGGCTTATGGGGAACATTTGGGGGACACATGACCATGGCTTTTATGACTCGGATTC	294					
Db	364 GGGGCTGGGGAACATATGGGGAGACCATGACACGCGCTTTTACGAGCCGGCCGC	423					
OY	295 CACTCTTTCGCTTCGATCACTCGTCTCAGATTCGGAATGATCCAAGATCTCTTCG	354					
Db	424 CGAGCTTCACATCTCGACACATCGCGCCGCCACAGATCCGATGATGAGAGTCCCTCG	483					
OY	355 CTTTCCCTCTGTTTGTGAGAGAGCTAGTA---AATGCCAAGAGATATAGTTGATGG	411					
Db	484 ATTCCCTTCCTCTCCGATACATACACGAGAGAAACCGAAGATATATAGTTGATGG	543					
OY	412 GTGTGGCATAGGTGCACCTCTAGATACCTGCGCCCAAGAAATTTGGAGCA-----CCAG	465					
Db	544 ATGTGCTATAGAGGCACTTTAGGTATCTGGCAAGAAATATGGGCTAATATTGTCTAG	603					
OY	466 TGTAGGCATCACTCGATCTCTGTTCAAGCTCAAGAACCAATCTCTGCTGCGGTCA	525					
Db	604 GGCATTAATCTCTCCACACCCCTGTGCAAGCCAGAGAGCTTAACAGCTTGCTGATGTCA	663					
OY	526 AGGATTTGCTCATTAAGGTTTCTTTCAAGTTGCTGACGCTCTACAGCAACCATCTCTGA	585					
Db	664 AGGATTAATATGGCAAGGTTTCTTTGAAGTTGTGATGCGTTGAACCAACCATTTCTGA	723					
OY	586 CGGCGCATTTGATCTGCTGTGTGCTCATATGAGAGATGAGAGCATATGCTTGACAAAGTAA	645					
Db	724 AGGGAAGTTGATCTGTGTTGGTGTGATGAGAGAGTGGAGAACACATCTGATTAAGAAA	783					
OY	646 GTTTGTTGAGAGATTAGCTCGGATGAGAGACCAAGCGCATTAATATATAGTAACATG	705					
Db	784 GTTTGTAATATAGCTGTGCTGTGCTCTCTCGGTGGAGAAATATATATCTGTTATAG	843					
OY	706 GTGCCACAGGATTTGGCCCTGACGAACATCTTACATTCATATGGAGCAAGATCTCTT	765					

DB	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	TITLE	JOURNAL	REFERENCE	AUTHORS
844	GI000000000	Arabidopsis thaliana	At1g64970	FL13011_27	mRNA, complete cds.	AY049258	1303 bp	linear	PLN 09-AUG-2001		
766	AA049258	Arabidopsis thaliana	At1g64970	FL13011_27	mRNA, complete cds.	AY049258	1303 bp	linear	PLN 09-AUG-2001		
904	AA049258	Arabidopsis thaliana	At1g64970	FL13011_27	mRNA, complete cds.	AY049258	1303 bp	linear	PLN 09-AUG-2001		
826	AT049258	Arabidopsis thaliana	At1g64970	FL13011_27	mRNA, complete cds.	AY049258	1303 bp	linear	PLN 09-AUG-2001		
964	AT049258	Arabidopsis thaliana	At1g64970	FL13011_27	mRNA, complete cds.	AY049258	1303 bp	linear	PLN 09-AUG-2001		
886	AT049258	Arabidopsis thaliana	At1g64970	FL13011_27	mRNA, complete cds.	AY049258	1303 bp	linear	PLN 09-AUG-2001		
1024	AT049258	Arabidopsis thaliana	At1g64970	FL13011_27	mRNA, complete cds.	AY049258	1303 bp	linear	PLN 09-AUG-2001		
946	AT049258	Arabidopsis thaliana	At1g64970	FL13011_27	mRNA, complete cds.	AY049258	1303 bp	linear	PLN 09-AUG-2001		
1084	AT049258	Arabidopsis thaliana	At1g64970	FL13011_27	mRNA, complete cds.	AY049258	1303 bp	linear	PLN 09-AUG-2001		
1006	AT049258	Arabidopsis thaliana	At1g64970	FL13011_27	mRNA, complete cds.	AY049258	1303 bp	linear	PLN 09-AUG-2001		
1144	AT049258	Arabidopsis thaliana	At1g64970	FL13011_27	mRNA, complete cds.	AY049258	1303 bp	linear	PLN 09-AUG-2001		
844	GI000000000	Arabidopsis thaliana	At1g64970	FL13011_27	mRNA, complete cds.	AY049258	1303 bp	linear	PLN 09-AUG-2001		
766	AA049258	Arabidopsis thaliana	At1g64970	FL13011_27	mRNA, complete cds.	AY049258	1303 bp	linear	PLN 09-AUG-2001		
904	AA049258	Arabidopsis thaliana	At1g64970	FL13011_27	mRNA, complete cds.	AY049258	1303 bp	linear	PLN 09-AUG-2001		
826	AT049258	Arabidopsis thaliana	At1g64970	FL13011_27	mRNA, complete cds.	AY049258	1303 bp	linear	PLN 09-AUG-2001		
964	AT049258	Arabidopsis thaliana	At1g64970	FL13011_27	mRNA, complete cds.	AY049258	1303 bp	linear	PLN 09-AUG-2001		
886	AT049258	Arabidopsis thaliana	At1g64970	FL13011_27	mRNA, complete cds.	AY049258	1303 bp	linear	PLN 09-AUG-2001		
1024	AT049258	Arabidopsis thaliana	At1g64970	FL13011_27	mRNA, complete cds.	AY049258	1303 bp	linear	PLN 09-AUG-2001		
946	AT049258	Arabidopsis thaliana	At1g64970	FL13011_27	mRNA, complete cds.	AY049258	1303 bp	linear	PLN 09-AUG-2001		
1084	AT049258	Arabidopsis thaliana	At1g64970	FL13011_27	mRNA, complete cds.	AY049258	1303 bp	linear	PLN 09-AUG-2001		
1006	AT049258	Arabidopsis thaliana	At1g64970	FL13011_27	mRNA, complete cds.	AY049258	1303 bp	linear	PLN 09-AUG-2001		
1144	AT049258	Arabidopsis thaliana	At1g64970	FL13011_27	mRNA, complete cds.	AY049258	1303 bp	linear	PLN 09-AUG-2001		

Tang, C.C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S.,
Davis, R.W., Theologis, A., and Ecker, J.R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
contributed equally to this work as PIs.

FEATURES

Location/Qualifiers

1..1303

/organism="Arabidopsis thaliana"

/db_xref="taxon:3702"

/chromosome="1"

/clone="RAF107-18-C14(R13225)"

/note="ecotype: Columbia"

1..67

5'UTR

68..1114

/note="gamma-tocopherol methyltransferase"

/codon_start=1

/product="At1g64970/F13011_27"

/protein_id="AAK83600.1"

/db_xref="GI:15146234"

/translation="MKATAPSSLSLTPYRTNSFGSKSLFRSPSSSSVSGMTT
RGNVAAPAAATSTALKKGLAEFYNETSGLMEETGDMHMGFPDPSSVOLSPGK
EAQIRMIESTLRAGVTDEEEKIKRVVDGCGIGSSRYLAKFGACIGITLSPV
QAKRANDLAAOSLAHKASFGVADALDQPEDGKFDLWSMESEHMDKAPKELY
RVAAPGRILITWCHRLMSAGEBALQPMENILDKICTKTYLPAMCSTDYVILLOS
HSIDGICADMSENVAPWPVAVIRATLTKWGLVSLRSKMSIKGALTMPLMEGKK
GVKEFGITCOOKPL"

3'UTR
BASE COUNT 352 a 267 c 307 g 377 t
ORIGIN
1115..1303

Query Match 37.4% Score 445.2 DB: 8 Length 1303;
Best Local Similarity 70.5% Pred. No. 3e-102;
Matches 626; Conservative 0; Mismatches 253; Indels 9; Gaps 2;

187 CAAGAGAGAGCTGAGAGAGGAGATGACAGATTTTACGACGATGCTGCTTATGGA 246
11
235 CACTGAGGCGCTAAGAGAGAGATGCGGATTTACATCAATGAACTTGGGTTTGGGA 294
11
247 GACATTTTGGGCGACACATGCGCTTTTATGCTGCGGATTCACGTTTGGCT 306
11
295 AGAGATTTGGGAGATCATATGCTATGCTTTTATGACCTGATCTTCTGCAACT 354
11
307 TTGGATCATGCT-----GCTGCTCAGATCGAATGATCCAAAGACTCTCTGCTTGC 360
11
355 TTCTGATTTCTGCTACAGAGAGCTCAGATCCGATGATTTGAGAGCTCTCTCGTTTGC 414
11
361 CTCTGTTTC---TGAGAGAGCTAGTAAATGGCCCAAGACTATAGTATGTTGGTGTG 417
11
415 CGGTGTACTGATGAAG 474
11
418 CATAGCTGACCTCTAGTACCTGCGCCAGAAATTTGGAGCAACCATGCTAGCATC 477
11
475 GATTTGAGAGAGCTCAAGATATCTTCTTAAATTTGGAGCAATGCTAGCTTAC 534
11
478 TGTAGTCTCTTCAAGCTCAAGAGCAATGCTCTGCTGCTGCTCAAGAGATTTGCTG 537
11
535 TCTCAGCCCTGTTCAGGCAAGAGAGCAATGATCTCGGGCTGCTCAATCTGCTCA 554
11
538 TAAGTTTCTCTTCAAGTTCTGAGCTCTTACAGCAACCATCTCTGAGCGCCAGTTTGA 597
11
595 TAAGCTTCTCTTCAAGTTCTGAGCTCTTACAGCAACCATCTCTGAGCGCCAGTTTGA 654
11
598 TCTGCTGCTCTCTTCAAGTTCTGAGCTCTTACAGCAACCATCTCTGAGCGCCAGTTTGA 657
11
655 TCTAGTCTCTTCAAGTTCTGAGCTCTTACAGCAACCATCTCTGAGCGCCAGTTTGA 714
11
658 TTTAGCTCTGCTCTTCAAGTTCTGAGCTCTTACAGCAACCATCTCTGAGCGCCAGTTTGA 717
11
715 GTTGGAGCTGTGCGCTCTCAGAGAGTGGATTAATTAATGATGATGATGATGATGATG 774
11
718 TCTTGGCCCTGAGCAATCTCTTACATCCATGAGAGCAAGATCTCTTAAAGAGATTTG 777
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11

Db 775 TCTATCTGCGGGGAGAGAGCTTTTGACGCCGTGGGAGCAAAACATTTGGACAAATCTG 834
Qy 778 CGATGATATTAACCTCCCTGCGCTGCTCAACTTCTGATTAATTAATTAATTAATTAATTA 837
Db 835 TAAGAGCTTCTATCTCCGCTTGGCTGCTCCAGCAATGATTAATTAATTAATTAATTA 894
Qy 838 CCTGTCACCTTCAAG 897
Db 895 CCATTCCTCCAGAGATTAATTAAGTTCGCGATTTGTCAGAGAGAGAGAGAGAGAGAGAG 954
Qy 898 AACAGTATACCTGACCTTACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 957
Db 955 TCGGCTATACAG 1014
Qy 958 AAAACGATTAAG 1017
Db 1015 GAAAGATTAATTAAG 1074
Qy 1018 AATTAAGTTTGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1065
Db 1075 CATTAAGTTTGGATCATCATCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1122

RESULT 3
AY090280 1047 bp mRNA linear PLN 24-MAR-2002
DEFINITION Arabidopsis thaliana At1g64970/F13011_27 mRNA, complete cds.
ACCESSION AY090280
VERSION AY090280.1 GI:19699149
KEYWORDS FLI_CDNA.
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 1047)
Kim, C.J., Chen, H., Cheuk, R., Meyers, M.C., Shinn, P., Banh, J.,
Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D.,
Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G.,
Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M.,
Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T.,
Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C.,
Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W.,
Theologis, A., and Ecker, J.R.

TITLE Arabidopsis ORF clones
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1047)
Kim, C.J., Chen, H., Cheuk, R., Meyers, M.C., Shinn, P., Banh, J.,
Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D.,
Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G.,
Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M.,
Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T.,
Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C.,
Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W.,
Theologis, A., and Ecker, J.R.

TITLE Direct Submission
JOURNAL Submitted (14-MAR-2002) Salk Institute Genomic Analysis Laboratory
(SIGAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA

COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the
sequencing and annotation of the RFLF cDNAs (RAF1 cDNA: 'RIKEN
Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y., and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the
sequencing and annotation of the RFLF cDNAs: Kim, C.J., Chen, H.,
Cheuk, R., Meyers, M.C., Shinn, P., Banh, J., Bowser, L., Chang, E.,
Dale, J.M., Goldsmith, A.D., Jones, T., Karlin-Neumann, G., Lam, B.,
Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J.,
Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C.,
Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A.,

gene	1. .1350	/db_xref="taxon:3702"	/organism="Arabidopsis thaliana"
CDS	/gene="g-TMT"	63. .1109	/gene="g-TMT"
	/codon_start=1	/product="gamma-tocopherol methyltransferase"	/protein_id="AA02882.1"
	/db_xref="GI:4106538"		
	/translation="MKATILAAPSLSLPYTNSSFSKSSLLFRSPSSSVSKTTT		
	RGNAVAALAAATSTELARKGIAEFYNETSGLMEELMGDMHNGFDPDSSVOLSDGHR		
	EAQIRMLIESLRFPAGVTDEEEKIKKVDGCGIGGSRSLAKFGACIGITLSLPG		
	QARKNALEAASRLSISHKRSFOVADALDOPFEDGKFDLVMSMESGFMHDKRFEVL		
	RVAPRGRRIITVYMCCHRILNSGEALDQPMENILDKICTKTYLPMSCSTDYVNLK		
	HSIDLIQCADMSENVAPFPMPAVITRALTMLKGLVLSLRGMSKIGALTMPLMBIRK		
	GVIFGIITCCPKP"		
BASE COUNT	374 a	278 c	314 g 384 t
Query Match	37.2%	Score 442;	DB 8; Length 1350;
Best Local Similarity	70.3%	Pred. No. 1.9e-101;	
Matches 624;	Conservative 0;	Mismatches 255;	Indels 9; Gaps 2
QY	187 CAAGAAGAAAGCGGCAAGGAGGACAGATTTTACGACGAGTCGTGCGCTTAATGGA	246	
DB	230 CACTGAGCGCGCTAAAGAAAGAAATACGGAGTCTTCAATGAACCTCGGGTTGTGGGA	289	
QY	247 GAAACATTTGGGGGCGACCAACATGCACCATGGCTTTTATAGTCGGATTCACGTGTTGCT	306	
DB	290 AGAGATTTGGGGAGATCATATGCAATCATCATGCTTTTATGACCCGTGATTTCTTGTCACAT	349	
QY	307 TTGCGATCATCGT-----GCTGCTCAGATCCGATGATCCAAAGATCTCTTCGCTTTC	360	
DB	350 TTCTCATTTCTGGCTCAACAAAGAGCGATGATCCGTATGATGAAGAGCTCTCCGTTTCCG	409	
QY	361 CTCGTCTTC---TGAAGACCGTATTAATAGGCCCAAGACGATATACGATGTTGGGCTGTG	417	
DB	410 CGGTTTACTGATGAGAGAGAGAGAGAAAAAGATAAAGAAAGTGTGATGTTGGCTGTGG	469	
QY	418 CATAGGTGGACGCTAGATACCTGGGCCAAGAAATTTGGAGCAACCAAGTGTAGCATCAC	477	
DB	470 GATTGGAGGAACCTCAAGATATCTTCTCTAAATTTGGAGCGATGATGATGGCATTTAC	529	
QY	478 TCTGATGCTGTTCAAGCTCAAGAGCAATGCTCTTGTGCTGCTCAAGATTTGGCTGA	537	
DB	530 TCTCAGCCCTGTTCAAGGCCAAGAGCAACCAATGATCTCGGGCTCAATCACTCTCTCA	589	
QY	538 TTAGGTTTCCCTTCAGTGTGCTGAGGCTCTACAGCAACCATCTCTGAGCGGCAGTTGA	597	
DB	590 TTAGGCTTCCCTTCCAAAGTTGGCGGATCTGTGGATACCAATTTGAAAGATGAAAAATTCGA	649	
QY	598 TCTGGTGTGCTCATGAGAGAGTGAAGCATATACCTGACCAAGCTAAAGTTTGTGGAGA	657	
DB	650 TCTAGTGTGTGATGAGAGTGTGATGATGCTGACCAAGGCCAAGTTTGTAAAGA	709	
QY	658 GTTAGTCTGGGTAGACAGCAACAGTGCCTATTATATATATGTAACATGTTGCCACAGGGA	717	
DB	710 GTTGTGAGCTGTGCGGCTCCAGAGAGTATGATATATATGATGACATGTTGCCATAGAAA	769	
QY	718 TCTTGGCCCTGACGAACAATCTTATATCATCCATGGGAGCAAGATCTCTTAAAGAAATTTG	777	
DB	770 TCTATCTGGGGGAGAGAACTTTTGAAGCGTGGAGCAAAACATCTTGGACAAATCTG	829	
QY	778 CGATGATATTTACCTCCGCTCGTGTGCTCAACTTTCGATTATGTTAAAGTGTGCTCAATC	837	
DB	830 TAAAGACGTTCTATCTCCCGGCTTGTGTCTCCACCGATGATTTATGTCAACTTGGTTCAATC	889	
QY	838 CCTGTCACTTCAGACATCAAGTCAGAGAATGGTCTCGCTTGTGTGCTCCATTTTGGCC	897	
DB	890 CCATCTCTCCAGATATTTAAAGTGTGGATGTGTCAGAGAACTAGCTCTCTTCTGCGCC	949	
QY	898 AGCAGTGAATACGCTCAACCTTTCACATGGAAGGGCTTATCTTCACTCTTGAGCAGTGGACA	957	

complete sequence.
 AC006193
 VERSION GI:4733953
 KEYWORDS HTG
 SOURCE Arabidopsis thaliana.
 ORGANISM Arabidopsis thaliana.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 118335)
 REFERENCE
 AUTHORS Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
 Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
 Gonzalez,A., Kremetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
 Lueros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V.,
 Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
 JOURNAL Unpublished
 2 (bases 1 to 118335)
 REFERENCE
 AUTHORS Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
 Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
 Gonzalez,A., Kremetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
 Lueros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V.,
 Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
 JOURNAL Direct Submission
 Submitted (09-DEC-1998) DNA Sequencing and Technology Center,
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA
 3 (bases 1 to 118335)
 REFERENCE
 AUTHORS Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
 Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
 Gonzalez,A., Kremetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
 Lueros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V.,
 Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
 JOURNAL Direct Submission
 Submitted (04-MAR-1999) DNA Sequencing and Technology Center,
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA
 4 (bases 1 to 118335)
 REFERENCE
 AUTHORS Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
 Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
 Gonzalez,A., Kremetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
 Lueros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V.,
 Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
 JOURNAL Direct Submission
 Submitted (11-JUN-1999) DNA Sequencing and Technology Center,
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA
 5 (bases 1 to 118335)
 REFERENCE
 AUTHORS Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
 Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
 Gonzalez,A., Kremetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
 Lueros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V.,
 Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
 JOURNAL Direct Submission
 Submitted (28-MAY-2000) DNA Sequencing and Technology Center,
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA
 On May 4, 1999 this sequence version replaced gi:4678192.
 COMMENT
 FEATURES
 source
 1..118335
 Location/Qualifiers
 /organism="Arabidopsis thaliana"
 /cultivar="Columbia"
 /db_xref="taxon:3702"
 /chromosome="I"
 /map="g1147"
 /clone="F13011"
 1092..2311
 /gene="F13011.1"
 join(1092..1225,1474..2311)
 /gene="F13011.1"
 /note="Hypothetical Protein"
 /codon_start=1
 /protein_id="A038245.1"
 /db_xref="GI:5042406"
 /translation="MGQIDTRKNTNTERNFTESTELSAFORYTAOITVMKIQO
 RRPTRKMKDKMQSSPETWTRIPDPDPSHTGVREDAMDSVCKPMKLYENPYCSCSSOS

OOHQHOKAFIWDLNFIVFMSESLGKADDEIKELKAEIDYERKARRAEMLIKLIK
 DVEEEMAREAEEMQNRKLEKELSSSEKSEVNRKRRLKEERQHRRLAEVLRERVOVK
 LMDARLEEKLESELEANEANGGERENRMKPKLIERACSSPARRRCENPQJRGINP
 FPRVMARISRSKSGSKLECKOVQGLILRQKTTTRCPRLSSPPP"
 complement(3352..3963)
 /gene="F13011.2"
 complement(3352..3963)
 /gene="F13011.2"
 /note="Hypothetical Protein"
 /codon_start=1
 /protein_id="A038246.1"
 /db_xref="GI:5042407"
 /translation="MGKCEGAGDEEDLLIKVKSQGVLEFSPVATGAFVSHGS
 GHAFSAVDLMLPLADHLIVGQSYLFPNVSELFTFVOSCHVRSSELSAT
 PYMSLDVNRVLRKSTYDVEFSNRSHIRTPHRLPHELFQKSTASVFG
 LLOIISDEGRTNELISVRVAKGFTSSITSSSENFSLSVQ"
 8107..9734
 /gene="F13011.3"
 join(8107..8140,8222..8358,8440..8489,8562..8893,
 8959..9062,9144..9219,9304..9734)
 /gene="F13011.3"
 /note="very similar to alcohol dehydrogenase; Similar to
 g111143388 class III ADH, glutathione-dependent
 formaldehyde dehydrogenase. [Arabidopsis thaliana];
 location of est 94M157 (g112757072) and 94M15X (3')
 (g113450565)"
 /codon_start=1
 /protein_id="A038247.1"
 /db_xref="GI:5042408"
 /translation="METQGVVIRKAAVAGAGEPLMEDVKVDPORLEVRIRILET
 SICHTDLSAKGGENARAPRILIGHAIVSVEGVEMAGDHVLEFCEGD
 CRCHRDGAKNGERFVDPKRVMTDGTGKTRFTSNKRIYELMTSFESEYVDS
 ACYKVDPLFPEKISLISGVSQDCIMEGGAAMNADIPASTVAIFGAVGL
 AVYKARAKASKITGIDINPDKFQUGREGISEFIPKESDAVHERNREIGVE
 YSECGASILEALEAFSLTNSGVGVTVMLVAGASPOLPHPELFQKSTASVFG
 FKETQLPEFTICLOGLMDLFLSIQPLPHDINEAMOLIHOGKALRCLHL"
 complement(10308..11958)
 /gene="F13011.4"
 complement(join(10308..10729,10802..10883,11042..11125,
 11208..11384,11461..11617,11723..11958))
 /gene="F13011.4"
 /note="Location of ests E6h377 (g112414136), 62B10T7
 (g11931264), 228115F7A (g111217083). 89% identical to
 membrane related protein CP5 [Arabidopsis thaliana]"
 /codon_start=1
 /product="Putative membrane related protein"
 /protein_id="A038248.1"
 /db_xref="GI:5042409"
 /translation="MMGLSLVFTEFHSGLFSLIVFAVPLMTAVAGVLCVWVWRPK
 WAYLDSKILNSPKPLNLOLPTSLKTSHHVNPSSSKETGTFVDDDERHMLKVE
 VKDGGCWIONMRSPTFSYQAMRDPENGGPOYSRTEFEDATPEMDVDFPMDFE
 RSKWMDMLYSTLERKDTGTWVQVNRKFPFCSQDREITIRRIDAGRYETCTIK
 GVQVPSVPQRNKRVDLYSSWCIRAVESKSDGEMTSCVLELFHEDGTFWEJAE
 LGVROGMGAVKRIEPLRAVQRAKAGALSPSAIMAHINTVSAEEFNNRSGIAE
 VTGDKPTGKNIPKILVVGAIALACTDKLTKRAVIFVARRFARMGKRM"
 13535..15717
 /gene="F13011.5"
 join(13535..13647,13808..14022,14535..14734,14807..153;
 15402..15717)
 /gene="F13011.5"
 /note="Identical to Tubulin Alpha-1 Chain [Arabidopsis
 thaliana] (g1135391) Location of ests 139J20T7
 (g112762968), 214C14T7 (g11159108)."
 /codon_start=1
 /product="alpha1 tubulin"
 /protein_id="A038249.1"
 /db_xref="GI:5042410"
 /translation="MRFLISIHIGAGIYGNSCMELCYCLEHGIQDPMTGSRSTVGA
 CHDAFNTFSESSESGHVPRAVFLDEPVIYIDVIRGTGOLPHRPLGSLLERLSD
 FARGHTVGREIYDTCLEKRLADNCTGCGIVPAVAGGCGSLGSLLERLSD
 FGKRSKLGFTIYSPQVSTAVEPYNSVSTSHLEHTDVVILLDNBAITDTCRSD
 IERPTYSNLRLLSQITSLSTSLRFGAIVNITVEQNLVVPYRTHPEMLSSYAVI
 SSAKAYHEDPSVEPITTSVPEPSNMMAKCDPRHGKGYACCLMARGDVPKDVNTAVAA

QY	544	-----TTCCCTTCAGGGTTGCTGACGCTCTTAAGCAACCATCTCTGACGGCCAGTTTCAT	598
Db	98364	TAGGCTTCCTCCCAAGTTGGCGGATGCCGTTGGATGATCGCATTCGGAAGATGGAAATTCGAT	98305
QY	599	CTGGTGTGCTCCATGAGAGAGTGGAGAGCATATGCTGACCAAAAGCTTAAGTTTGTGGAGAG	658
Db	98304	CTAGTGTGTGATGAGAGAGTGGAGCATATGCTGACCAAAAGCCAAATTTGTGTAAGAG	98245
QY	659	TTAGCTGGGTGAGGACGACCAAGTCCATTTTAAATATGATGAACATGGTGGCACAGGGAT	718
Db	98244	TTGGTACCTGTGGCGCGCTCCACGAGAGTGGATATATATATGATGATGTTGCTCATGAAAT	98185
QY	719	CTTGCCCTCGACGAACAATCTTACATCCATGAGAGACAAGATCTCTTAAGAAGATTTGC	778
Db	98184	CTATCTGGGGGGGAGGAAGCTTTGACAGCCGTTGGAGACAACATCTTGGACAAAATCTGT	98125
QY	779	GATGATATTTACCTTCCTGCTGCTGCTGCTCAACTTCTGATATATGTTAAGTTGCTCCATTC	838
Db	98124	AAGAGTCTATCTCCCGGCTGGTGCTCCACCATGATGATATGATCAACTTGTCTCAATCC	98065
QY	839	C-----	839
Db	98064	CATTCCTCCAGGATATATATTCCTCCATCCCTTGTGCTGCCAAACATTTACTTAATAG	98005
QY	840	-----TGTCACTTCAGACATCAAGTCAGAAAGTTG	870
Db	98004	GTTTTTTACTAACTAGAACTGAAATATGTTTTTATGATATTAAGATGTGGCGGATG	97945
QY	871	GTCCTGCTTGTGTGCTCATTTTGGCCAGCATGATGATGATGATGATGATGATGATGATG	930
Db	97944	GTCAGAGAACGTAGTCTCTTCTGTGCGGCTGATATGATGATGATGATGATGATGATGATG	97885
QY	931	TCTATCTTCACTCTTGGACGACGTGACAAAAAAGATTAAGAGGCTTTGGCTATGCTCAT	990
Db	97884	CCTTGTGCTCTGCTCTTGAAGGATGATGATGATGATGATGATGATGATGATGATGATG	97825
QY	991	GATGATTAAGGAGATCAAGAAAGATCTAATTAAGTTTGGCATCATTCATGCAAAACC	1050
Db	97824	GATGATTAAGGAGATCAAGAAAGATGATTAAGTTTGGCATCATTCATGCAAAACC	97765
QY	1051	TGAATTAATGGAGAG	1065
Db	97764	ACTCTAAGCTTAAGAG	97750
RESULT 8			
AP003587		43550 bp DNA linear	BCT 28-NOV-2001
LOCUS	AP003587		
DEFINITION	Nostoc sp. PCC 7120 DNA, complete genome, section 7/19.		
ACCESSION	AP003587	BA000019	
VERSION	AP003587.1	GI:17130808	
KEYWORDS			
SOURCE	Nostoc sp. PCC 7120 DNA.		
ORGANISM	Nostoc sp. PCC 7120		
REFERENCE	Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.		
AUTHORS	1		
	Kaneko, T., Nakamura, Y., Wolk, C.P., Kuritz, T., Sasamoto, S.,		
	Watanabe, A., Iriuguchi, M., Ishikawa, A., Kawashima, K., Kimura, T.,		
	Kishida, Y., Kohara, M., Matsumoto, M., Matsuno, A., Muraiki, A.,		
	Nakazaki, N., Shimpo, S., Sugimoto, M., Takazawa, M., Yamada, M.,		
	Yasuda, M. and Tabata, S.		
TITLE	Complete genomic sequence of the filamentous nitrogen-fixing		
JOURNAL	Cyanobacterium Anabaena sp. strain PCC 7120		
MEDLINE	DNA Res. 8 (5), 205-213 (2001)		
REFERENCE	21595285		
AUTHORS	2 (bases 1 to 343550)		
TITLE	Kaneko, T.		
JOURNAL	Direct Submission		
	Submitted (02-MAY-2001) Takakazu Kaneko, Kazusa DNA Research		
	Institute, The First Laboratory for Plant Gene Research, Yana		
	1532-3, Kisarazu, Chiba 292-0812, Japan		
	(E-mail: kaneko@kazusa.or.jp,		
	URL: http://www.kazusa.or.jp/cyanobase/,		

Tel: 81-438-52-3935(ex.2338), Fax: 81-438-52-3934)

```

FEATURES
    source
        Location/Qualifiers
            1..34350
                /organism="Nostoc sp. PCC 7120"
                /db_xref="taxon:103690"
                /note="synonym: Anabaena sp. PCC7120"
                complement(94..411)
                /gene="a111719"
                complement(94..411)
                /note="a111719"
                /gene="a111719"
                /note="ORF_ID:a111719"
                hypothetical protein"
                /codon_start=1
                /transl_table=11
                /protein_id="BAB73418.1"
                /db_xref="GI:17130809"
                /translation="MRKKSAILEAVHETADLHKAGLMNQTTLREPHLCPIEPL
                EPLDIKEIRESSQVAVFAKILNISTSTVOKWEIGOKRPSGASLKLHLVKNRGLNS
                VLY"
                complement(504..782)
                /gene="a111720"
                complement(504..782)
                /note="a111720"
                /gene="a111720"
                /note="ORF_ID:a111720"
                hypothetical protein"
                /codon_start=1
                /transl_table=11
                /protein_id="BAB73419.1"
                /db_xref="GI:17130810"
                /translation="MKIKNRTFDRMARKEGIKNLISICNAVENAPGLYDAGGGLF
                KRIKAPGKSGSGFRLLIATNNEVGLFLAFQKTNVTLTHMKKL"
                935..1507
                /gene="a111721"
                935..1507
                /note="a111721"
                /note="ORF_ID:a111721"
                /codon_start=1
                /transl_table=11
                /product="transcriptional regulator"
                /protein_id="BAB73420.1"
                /db_xref="GI:17130811"
                /translation="WRHKDKNKQAIICDAIEILTANGPADTSMKIAKANVSPATI
                VYFENKREKLNTKLYLFVKQEISAEILGVNQNLSEKAFKIMNNYKVAIKNVRF
                AFTEQFANSPWDRICKDESLSYKPMDLNRRGKEEVFDIPLEIFTAFFAVLTG
                LIKEHFGALIDELKMLMETTYKIAMDAVTN"
                1569..2504
                /gene="a111722"
                1569..2504
                /note="a111722"
                /note="ORF_ID:a111722"
                probable oxidoreductase"
                /codon_start=1
                /transl_table=11
                /protein_id="BAB73421.1"
                /db_xref="GI:17130812"
                /translation="MKNEKNAENIPPOKGRLATYVSSSGIGYETARLANKQASVI
                IAVNRIDKGNKALAKILIOONKADADYKLMELDIANASVKNFAENRKNYLRIDLLINN
                AGVNIPTSKTDEPELOFGNHLGHFALTOGLEFLSTESRIVNVSSGHNHKKI
                DPDDLMEORSYAKKAYGDSKLANLYTYVELDRKLKNGIDTLVTASHPGWTATELO
                RTAGGIVATLNGIVADITMGALPTLRALIEGLGAGYFGPGNFMERGYPIKVESN
                ELSKDQALAKLWVSEKLTGVKEFNKKAOSAGK"
                complement(2701..3462)
                /gene="a111723"
                complement(2701..3462)
                /note="a111723"
                /note="ORF_ID:a111723"
                probable aldolase"
                /codon_start=1
                /transl_table=11
                /protein_id="BAB73422.1"
                /db_xref="GI:17130813"
                /translation="MSVFSRPLPVPFVNSIEEERLHKORLAAAFRLFGKPGFSEGIAG
                HITADPEFTDHFVWNPILGKTYGHIRVSDVLVYNRDGVDVKGDALINQAAFAIHSOIH

```

```

        EARPVVAANAHSIYCKAMSSGLRILPILTDACAFEDHALLDDYTGCVLETSCK
        RLAEFTLGNKKALILRNHGILTVGHTVDEAAFWYISLERSCOQLLAAGOKPILIGHE
        TARLTHSQVGEPPKAGWSEFQPLVDRAAREPELFD"
        complement(3764..5167)
        /gene="a111724"
        complement(3764..5167)
        /note="a111724"
        /note="ORF_ID:a111724"
        unknown protein"
        /codon_start=1
        /transl_table=11
        /protein_id="BAB73423.1"
        /db_xref="GI:17130814"
        /translation="MSSTSQGVGGLPDLTHPELILQGSQILITASLLVFLVIALCIA
        IALMSFALHNSQOQITFIGEWAIRYAOILRGLHIALVILLVGGFFLSTLSNRH
        HWEAKIAQVQVSGDKLEOIAPOVRYATQEPYRNQGNKIKVANDTOOVSRLS
        LAGSOIQKLESTDVQGSAYRARDYATPQYVNRKIDINNEFPELOPNKYSILSN
        KVERDRTLEFVNFGNVSFAERLEGGQSTFRVYTOAGGCRWYNSAGQILSNRFL
        TTIANFGADFSAGIVPKERKADGRSTQYTWLFDNVSYNKPFQVFTNPIRQGTII
        PRLLAPALFLMWILLILYLSLPISTVNAIASISYIFACILTLTLYLSRMDAQLAMTM
        ISLVLETLTWGSGSTRSASIAIICTIAGAVVPVGLVLPVSGLTSLVAGLLSVILVA
        VRHMYNMNIWQPEDSKS"
        complement(5492..5902)
        /gene="a111725"
        complement(5492..5902)
        /note="a111725"
        /note="ORF_ID:a111725"
        unknown protein"
        /codon_start=1
        /transl_table=11
        /protein_id="BAB73424.1"
        /db_xref="GI:17130815"
        /translation="MELKPSGLMWJASLTUNGIAITYSGRSAIAEAKSNFYLQVA
        DQGRPIADLASINGTNTFTIQQGEGIAKRRILAAACQGLFTRKKSLYANAKRLT
        SEDKTBQETFEPRKPTPDASMETVCSVANNRQK"
        6321..6842
        /gene="a111726"
        6321..6842
        /note="a111726"
        /note="ORF_ID:a111726"
        /codon_start=1
        /transl_table=11
        /product="transposase"
        /protein_id="BAB73425.1"
        /db_xref="GI:17130816"
        /translation="MKSQHLQATTLAGEVSOQYIKETALNIAELQDFIDMCPDARE
        VKKALAVKLYOGYILDEYQKILDSRSSTGKKAKEETGIDGLRANKGGRKSYLNA
        KQLOEVLVSWLQTRDCWELGELETOLAFETDVVYESKOSTYDLFSEAGISMKRTKYNP
        KADENAVAEKKRD"
        6919..7431
        /gene="a111727"
        6919..7431
        /note="a111727"
        /note="ORF_ID:a111727"
        /codon_start=1
        /transl_table=11
        /product="transposase"
        /protein_id="BAB73426.1"
        /db_xref="GI:17130817"
        /translation="MWGDISGYWKGKSDOEIAVPVNERKQTYGADVILIGELVLYK
        AYDAGSKQNTINYLEYLANSPDORLLIWDGASYSRSEIRGFDSVQOSLPTEDWK
        IHCVRAPNCPVQNPREDIWLQAKTVVRRFCALIPFSHLKIFEFWFIHRTFDEDTL
        OMVGYTSKIR"
        8064..9401
        /gene="a111728"
        8064..9401
        /note="a111728"
        /note="ORF_ID:a111728"
        hypothetical protein"
        /codon_start=1
        /transl_table=11
        /protein_id="BAB73427.1"
        /db_xref="GI:17130818"

```



```
/gene="slr0073"
/note="ORF ID:slr0073"
/codon_start=1
/transl_table=11
/product="sensory transduction histidine kinase"
/protein_id="BA10540.1"
/db_xref="GI:1001703"
/translation="MGEMPSFEGTTLDOOKILGYFIEEAQHELETLENGIILDGKY
VQSEYVEMERASHSTKGAAMGYSTIOKTARLEDAEKVLKEHOLPNQKLESLF
LNGIDVLQDLVEKIQSPAGIQPEPADSTVDAQIQLPELQNYLWLLNQSATRTSTQ
TMAKSTSTGTFENESDVPEETKVLHMLDLFKQPSATSEKLEQYCTRLGRILAPKO
ENMOHLSTLSKQALANPLHSHRTLAPVLLDKQGDQFVLAKPGNIRVSDGLQKLS
SOVAOVLIPADPAASVLMKVFENDSOLROLIKMLSAAR"
complement(1930..2652)
/gene="sl10088"
complement(1930..2652)
/note="ORF ID:sl10088"
/known_protein="
/codon_start=1
/transl_table=11
/protein_id="BA10541.1"
/db_xref="GI:1001704"
/translation="MGVLSYVWTATRRFONFLLLTMTLSSSHSTKEDILCYLLKE
GQSAIEMAEELIGSPQAMRKHLKDLKLENDGLIEHQEROSMGKRPQFLYKLSKQREOF
PQRYGEFALSPIDSLVEVGEEOGLAVLKQKQKAEAYRQOIGQPLSKVHKLVEL
RROGVMAEIRHPLSVBOAEKFTLSEHNICALDVASSTPYGCHLEMAALLPQCAIE
RTHMLNDGEHTCGYLIOKTPN"
2908..4350
/gene="ycf24"
complement(4350..4350)
/note="ORF ID:slr0074"
/codon_start=1
/transl_table=11
/product="ABC transporter subunit"
/protein_id="BA10542.1"
/db_xref="GI:1001705"
/translation="MSSITVKNLYNQPKGVFNITENADIPRGLSENVVLLSKKN
EPEMLDRLRYRHMILMAEPTRPAHYHPIDODITTSAPKQSKKLSLSDVDP
ALLTEPEKLGIPLEOKRLSNVAVADAFDSVIGTTEKEKLAEDGVIFCSISEALQEH
PDLQKYLGVVPPADNFPALNSAVDSGSEVFIPKGVKCPMELSTYERJNNGDQO
FERLLIAEBSGASVYLEGCTAPMDTQNLHAAVEIALDNADIKYSTVQNTVAGDE
NGKGIYFVTRKGLKGVNSKISMTQVETGSAITWKYPCGVLDGNSNGEYSIALT
NKKOQADPTGKTHIHGKTRBSIITSKISAGNSANSYGLYKMGPKAGANYSQDGS
MLIGDRAANTFPIYQVNDNTAKVHEHSGTSGEDQLFTRQKISSEDAVSNLVGS
FCKDVLNLEPMEFAEADKLSTLKEGTG"
4492..5262
/gene="ycf16"
4492..5262
/note="ORF ID:slr0075"
/codon_start=1
/transl_table=11
/product="ABC transporter subunit"
/protein_id="BA10543.1"
/db_xref="GI:1001706"
/translation="WSOTILSKNLTASVDGNQILKGVNLEIKAGEVHAIMGRNSGK
STLSKVITGHDPDETGTGEIILYQGDLSALPHERALAGIFLAQYPLEIPGVSNLDF
LRIAYNARKHLGLELDPDFEDLQEKLDVYKMPALFERSLNEGSGGKKRNEI
LQMIAPETLSLDEIDSGLDIDALRIYSEGVNFKNPDNATLVITTHQRLNLTITPD
HIHYMDGKIVMSGKELALEEKGYFLDEQVLAFL"
5398..6759
/gene="slr0076"
5398..6759
/note="ORF ID:slr0076"
/hypothetical_protein="
/codon_start=1
/transl_table=11
/protein_id="BA10544.1"
/db_xref="GI:1001707"
/translation="WTAIADALQKVHQTREIDKEALKEQAMVSGTEDSILVNRK
```

```
gene
CDS
/gene="nifs"
6827..8089
6827..8089
/note="ORF ID:slr0077"
/codon_start=1
/transl_table=11
/product="nifs protein"
/protein_id="BA10545.1"
/db_xref="GI:1001708"
/translation="WVALQIPSLATVRQDEPILNQEINGHPLVYLQNAATSQRPAY
LEKLMHYENDNANVHGAHOLSVRATDAYEAVRNKVKAFINARSFRELIVRNATTEA
INLVASGMNNLKAGDEITITVMEHNSLVPMQWVAKTGAVLKQFVQDEQSFLE
HKPTLSSEKTKLVYVYHISNMLGVNPAEETLAOLAHQAGARVLDAQOSAPHYPLDVO
LIDCDWLVASGKRCAPPTGICFPLYKEKELIEMAPFFEGGEMIAVEFDDHTTGLPH
KFEAGTPAIAEIALGAADVTLTDGEMENINHYVELTHYLMQGLQIPQLKLGPNP
KHGDAALASFPVAGILHASDVATWVDQGIARSGHCTOPLHRLFLDASGSARASLYP
YNTKEIDLEFLOSQATIRFESDDDFV"
complement(8079..9269)
/gene="sl10086"
complement(8079..9269)
/note="ORF ID:sl10086"
/hypothetical_protein="
/codon_start=1
/transl_table=11
/protein_id="BA10546.1"
/db_xref="GI:1001709"
/translation="MRVILMTGKGVGKTSVAATGLRCACELGKTLVLSPPAHSLA
LSFDELEIHEPRLYKEMGAEIDLMELEGMGAVKRYITQVQARGLDQVQAEELA
IDGMDLEIFGEPRYRMRKRYHDEADVLLIDSAFTGALRLSLPEVGGVYMRFEYPLQ
GNSVALRPLVEPLFRPIGSESLPKDEVMDAPVEFEQJLEALEKYLTDMTQSVRLVTN
PEKMLKESLRHAYLSLNVSTDLVIANRILPEYIDDPFQRMKSNQOYKQELTYDN
FHLPLVVKAPLESEEMCGALAEKRLKDTLYKDEDSQVYKENTINIVQGSNDLSLE
AVKR"
complement(9312..10523)
/gene="sl10085"
complement(9312..10523)
/note="ORF ID:sl10085"
/known_protein="
/codon_start=1
/transl_table=11
/protein_id="BA10547.1"
/db_xref="GI:1001710"
/translation="MNSDHSRQPLFSLIPHLMGKGHIVYGAVRSVYELGMQY
QAMVQDADPNLEEGDRLHRIILPEDELEAESMAKITYRLQAMGCSVLAQQLSL
LNGESRPVITLERPIHOLFAIMWGRQOLSPPARSRIHILYRVDVQVLTSTSY
HMLHQRKLLPPLPQCFITPDEPLGRSLQDYFCQPTVMYPRHDEAVLQVLEFAD
PILCWMPGAPRERKGSVNRHLVSLRPNPSPLDLAASGLKDSBESIEYHLIGDR
LITREBAGMLANKAOIILLVYDEAAVARTSGIFTEAIIAERIPLVSPGTWARRLEINQY
LDSIVVDQAOOSLVAEIDQRLADSSIHQOLAAAOAVRRHSHSLESYAKTLOETWLS
RINGNNEVNNM"
complement(10520..11005)
/gene="nifs"
complement(10520..11005)
/note="ORF ID:sl10084"
/codon_start=1
/transl_table=11
/product="imidazoleglycerol-phosphate dehydratase;
histidinol-phosphatase"
/protein_id="BA10548.1"
/db_xref="GI:1001711"
```


REFERENCE
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoidae; Andropogoneae; Sorghum.
 1 (bases 1 to 108553)
 Song, R., Liaca, Y. and Messing, J.
 Mosaic Organization of Orthologous Sequences in Grass Genomes
 JOURNAL
 Unpublished
 2 (bases 1 to 108553)
 Song, R. and Messing, J.
 Direct Submission
 TITLE
 Submitted (08-JUL-2002) The Plant Genome Initiative at Rutgers - Waksman Institute, Rutgers University, 190 Frelinghuysen Road, Piscataway, NJ 08873, USA
 JOURNAL
 Location/Qualifiers
 1. 108553
 /organism="Sorghum bicolor"
 /cultivar="Brix623"
 /db_xref="taxon:4558"
 /chromosome="G"
 /clone="BAC SB_BBC0234M12"
 /note="overlaps with sorghum BAC SB_BBC0126P21"
 1. 108553
 /note="pnp200725 orthologous region"
 complement(<4347..>7716)
 /gene="SB234M12.1"
 /note="similar to Triticum aestivum Yr10 gene
 GB:AF149112.1 GI:11990496; similar to Hordeum vulgare
 M1a1 gene GB:AY009939 GI:11612212
 /evidence=not_experimental
 complement(join(<4347..4451,4587..4676,4917..4990,
 5038..6794,6896..>7716))
 /gene="SB234M12.1"
 /product="putative stripe rust resistance protein Yr10"
 complement(join(4347..4451,4587..4676,4917..4990,
 5038..6794,6896..>7716))
 /note="SB234M12.1"
 /note="putative 108 kDa protein"
 /codon_start=1
 /evidence=not_experimental
 /product="putative stripe rust resistance protein Yr10"
 /protein_id="AA094319.1"
 /db_xref="GI:22208504"
 /translation="MEVTVGLTGLVFKLGLVGEYNLQGVKEIMFLQELQSMQ
 GALKESKAPSELDHODRIMASEVRELSYDIEDSIMFVRCGGOOLLAAPGMSK
 FIDRSIDLTRFRARROVAMDIRIKRVI EARRRRAVYKIDGVGAPVDVDRLLA
 HYRAVPLVGTIDARDELIVTDDGSEASKGRVSYVCGGIGKTTIANYVOK
 IRFDPCMAFVSQTPDMRLEBGLISELGDINESTRVRFIDAIKFLQTRKC
 IVIDDIDISVWKIRCALPDNMGYVITTRNFKVAEEIGAVSKALKCHESSRLV
 FYRIFGNEKRYCPDEHLTEVSHRINKCAGVPLAITASLANKARDMEWLEVY
 NSIGTGLEDSDVENMRKILAYSYDLKYLRVCLATLSMPEPDYPIYKHLIMMILA
 EGFVCGEGKSLPELCFCEYENELINTSIOPVYVHRHAMI EHCYVHMVLEVISLN
 EDNFVITLLNBSHSSOSKMRRLSDNSVYHLASPPSTSMQVRSYITSSALDQ
 PALASFRVLVNLKGSYLPOGCDLKHGNLEFLRYGLGRTYNNELPEETGNIAYLO
 TLDVVGYSYLSSTVQVLRHLMCLVDONTVPNGIGRLTALELSTLYLQCDSTDI
 PEECLHLETRVLELFPMNNTLEKSLVACLKLRKIOSIVIMASGGEVNFAMAPRH
 VORLOJOSCFEPLDPMNPSPLGLSFLIMVRLROEDLETLGRPLRALRYLKLKH
 EKLKIPRFTVAGCSPFLCGLGEGVAFERHOGAMITKRLNLTLYADVADEITNA
 SSDGVOCYIITFRSGANMEVEKAKAAYASTTEITYETLNTLQOQKKKMMAMCKE
 IHIDELGRRNMGIOTRSGKRSNRYVPLLRCK"
 8776..8802
 /rpt_unit=gga
 complement(<12094..>14263)
 /gene="SB234M12.3"
 /note="highly similar to Oryza sativa gene for putative
 receptor kinase GB:AF128457 GI:4680335; similar to Oryza
 sativa receptor kinase-like protein gene GB:AF119222
 GI:4680488; similar to Arabidopsis thaliana receptor-like
 protein kinase emb:AL138654 GI:6899879"
 /evidence=not_experimental
 complement(join(<12094..12725,12823..>14263))
 /gene="SB234M12.3"
 /product="putative receptor kinase"
 complement(join(12094..12725,12823..14263))

/gene="SB234M12.3"
 /codon_start=1
 /evidence=not_experimental
 /product="putative receptor kinase"
 /protein_id="AA094320.1"
 /db_xref="GI:22208505"
 /translation="MVTLLARPLRPLFLLLAAAAAANGNMTAEAAALVNL
 KRSFTDPLGRLEAMSATPPPCDAASPVCQCKGSLVGRILHMLSGTFDGA
 AKLPLRHSVNLKHNALSGPLPASLGTLRGLRALYLSNNFSGPIPADVEANRMLKTL
 YLDNNRITGPIPADAIAPRLMEHLDRNIDCPVPKLPESLKRFRVNSHRLSGTL
 PDSVAERTDASSFRNGELCGSPGSDAIVCYAGALPAPPPATPADYNAAMEETSV
 FVAIGIILLVILVGAVALMLRDERNSVATMDYTAGSTTGAGAGAGAGSSKSLVA
 DASTTAAPRAGAVAVDMASSGSSSHGCGSSSGRMRGEFVLLNMDIAPGLPDL
 MKASAEVNGCTLSGAYKAARNGVYAVKRLRPMNRGRDEPQHVHMLGQLRHPV
 LPVGYHRRKEEKLIVSEYMPRGSILYTLHGDSNRYVILDMOGRVRYAVGVNGLAV
 LHERGIGTAGRLVSMDCGDPAPPPRPHGMLKSGNILLGADMERLYVAGFPPLVNA
 AADPAMPAFRNSPBGTTGAVSARSADVCLAVLELTGTFRPSQYLLSAGGTDVH
 WAATAVAGGAEALVDPALIAAGGDAVRLRLVGHGASPEECPSPVAEAMVVEET
 GSAAGAGS"
 12558..12578
 /rpt_unit=cgg
 complement(13145..13307)
 /rpt_unit=cgg
 13490..13517
 /note="GC-rich"
 14190..14216
 /rpt_unit=cgg
 complement(<15173..>20442)
 /gene="SB234M12.4"
 /note="similar to Arabidopsis thaliana gene for putative
 sugar transporter protein GB:AY094465 GI:20453188"
 /evidence=not_experimental
 complement(join(<15173..15605,16042..16226,16382..17055,
 17854..18423,18531..18650,19447..19501,19596..19649,
 20293..>20442))
 /gene="SB234M12.4"
 /product="putative sugar transporter"
 complement(join(15173..15605,16042..16226,16382..17055,
 17854..18423,18531..18650,19447..19501,19596..19649,
 20293..20442))
 /note="SB234M12.4"
 /note="ATG435300/F23E12_140"
 /codon_start=1
 /evidence=not_experimental
 /product="putative sugar transporter"
 /protein_id="AA094321.1"
 /db_xref="GI:22208506"
 /translation="MSMTHCKLNTVOAEDLMGATGACQGVOMCGKLEAVAVLQCAH
 RIMAEEVSTVHQEVCKMTFMPGAAIASRLIDCEAYLRTFESGFIYCKRAIVLDEM
 STVEFAVVVSIGVAMLGMDPALLEANHMKKEFLNGPSLEGITTLAASFGAIVIT
 IFSGALLDMLRRALIVSSIVLPSCGYLMSRPITYVILLARLIVSGSGLVPFCV
 IYISTSPNNRNGGLGTAPQEFRTGTFSTCLIFMLTMSPPMRHITGATIPASIV
 YFALLVYLSGSPRWLASDGIISEARVSLRLRCKRHVSGELAVTVGVYISDSAV
 GTARALVSPGSGASATWPSRSTFYWLSDPLVLDLSIHNNSEGSRRNSFEPVENS
 SEPEHEHNHRSDQDOTTEAYAGCVNNGDGLRASLLQAASEVNDANTSTSE
 GSSVFLRHGTSYLAOEMASIHDDIEEERIEHGFVPHOSAPDMSRRHOVRL
 SETDAMKRRKVVILQPIRLALCTOMLLOAQOSGTGILRCAPETILEOVGSLFSD
 IGLSHSTSIILSTLOALLILPCTAAMLMDVCGRRVLYATTPITLISLVSMSKI
 LFNKQFRAIVFHALVFCSTVYIGAPYIPNLICSSIFPTKARATCASCSLSFW
 GGLLSAYCLPVMWSTYIGAGAGIYALCCAPLFLFYRIBETKMLINLEIALEFLKS
 ROEYVQ"
 complement(22125..22161)
 /note="AT-rich"
 complement(22424..22446)
 /note="AT-rich"
 23019..23089
 /note="AT-rich"
 23130..23172
 /rpt_unit=ta
 23495..23526
 /note="AT-rich"
 26812..26900
 /note="AT-rich"

```
repeat_region      37014..37038
                    /note="GC-rich"
repeat_region      complement(37399..37462)
                    /rpt_unit="a
gene               <37578..>44288
                    /gene="SB234M12.6"
                    /note="similar to Nicotiana plumbaginifolia up1 gene
emb: A272011 GI:6996559"
evidence="not_experimental
join(<37578..37721,38745..38806,38998..39100,40241..40317,
40518..40579,40690..40777,40913..40968,41080..41110,
41205..41318,42137..42222,42364..42537,43066..43258,
43335..43537,43864..43887,43980..>44288)
/gene="SB234M12.6"
/product="putative oligouridylylate binding protein"
join(37578..37721,38745..38806,38998..39100,40241..40317,
40518..40579,40690..40777,40913..40968,41080..41110,
41205..41318,42137..42222,42364..42537,43066..43258,
43335..43537,43864..43887,43980..44288)
/gene="SB234M12.6"
/codon_start=1
/evidence="not_experimental
/product="putative oligouridylylate binding protein"
/protein_id="FAM94322.1"
/db_xref="GI:22208507"
/translation="MAPQRKKAQAQOOLMOALLDQQAQPPPEFGHHHPGLL
AAPQMPPIVSGNLPQFDSSTCRSVVGNLSLVTDVLDVEFQSIGPVGCKLIRE
KSFDFDIYDRYALAILSLNGRPYGPPIVMAVYSTOREDSGNHINFEVDLC
PEVTDATLFAFGSYGTSQDARVMDQKTRSGFGFVSFRNOQDQASANDLNGKL
GNROIICNNATKAGNAGEEKOIYDSKVDLITNGSESKENEDGPNOPFTVYVG
NLPHETNNNDVHLFPHSISAGSIEVRVTPDKGFPVRSSTHEAQAQLOMANOILG
GROIKTTIYNGYKTLRAVESDFCEVLIVSRVTLVASLVYVASAFLEGNDAVQ
GIDLTAASTAGAVGCGVLGVLTVTLVVAOSPEGLITPMKGAACVATYVR
SKLGGVKECAKGEAVVAVRVTENLGMCSQSPKPPPTASPLPPAPALFSTGV
SPTDFLAIOMALNKMATNPISLMGHALNOVALGMGRASOSIYDGGFOGISTVAATT
SAQOQOQOQOOLMY"
repeat_region      37580..37668
                    /rpt_unit="caag
                    complement(39121..39142)
                    /note="AT-rich"
repeat_region      44247..44274
                    /rpt_unit="caag
                    complement(45252..45289)
repeat_region      /rpt_unit="cg9
                    complement(45290..>46757
                    /gene="SB234M12.7"
gene
Query Match      8.4% Score 99.4; DB 8; Length 108553;
Best Local Similarity 71.8%; Pred. No. 3.9e-14;
Matches 130; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
OY 367 TTCTGAGAGCGTAGTAATGCGCCCAAGATAGTAGTGGTGGCATAAGGTGG 426
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 94217 TGCAGATGATCTAGAGAGAGCGCCCAAAACAATGATGACGTAGATGTGGAGG 94158
OY 427 CAGCTAGATCTGCGCCCAAGAAATTGGAGCAACAGTATGACATCATCTGAGTCC 486
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 94157 CAGCTCAAGGTACTTGGCCCAAGAAATACGAGACAAATGCGGACATGAGTGAAGCC 94098
OY 487 TGTTCAAGCTCAAGAGCAAAATGCTTGTGCTGCTCAAGATGAGTGGTGAAGGTTTC 546
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 94097 TGTTCAGAGCTGAGAGAGAAATGCTTACTGCGCCAGGCTTGTGAGACAGGTGCC 94038
OY 547 C 547
DB 94037 C 94037
```

```
RESULT 12
AX073663      930 bp      DNA      linear      PAT 06-FEB-2001
LOCUS
DEFINITION      Sequence 7 from Patent W00104330.
ACCESSION      AX073663
```

```
VERSION      AX073663.1 GI:12710084
KEYWORDS
SOURCE      Synecocystis sp. PCC 6803.
ORGANISM      Synecocystis sp. PCC 6803
REFERENCE      Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
AUTHORS      1 (bases 1 to 930)
TITLE      Herbers, K., Badur, R., Kunze, I. and Geiger, M.
JOURNAL      Identification and overexpression of a dna sequence coding for
FEATURES      2-methyl-6-phytylhydroquinone-methyltransferase in plants
source      Patent: WO 0104330-A 7 18-JAN-2001;
              Sungen GmbH & Co. KGaA (DE)
              Location/Qualifiers
              1..930
              /organism="Synecocystis sp. PCC 6803"
              /db_xref="taxon:1148"
              10..915
              /note="unnamed protein product"
              /codon_start=1
              /transl_table=11
              /protein_id="CAC8417.1"
              /db_xref="GI:12710085"
              /translation="MAAGLYLTLARGYQSSDVAMAYDQNTEDGILEYWGDIHLGH
              YGDPVAKDFIQSKIDFVHMAQMGIDLPPTGVYLDVCGIGSGSRILAKDGFNV
              TGITISPOOVKRAVELTPDVTKAFVADAMALSPDGSDVYMSVACGPHMDKAF
              AKELLRVKRGGITLVYADNRDROVPLFWFKPVYRROLIDQMSHPAFASIGPAEN
              LEATGLVEGVTTRADTVPLPLPHMDITIMOGITIRPOGILDTGIRGFTKSVREPTILL
              MRLAFGVGLRFPFMRVARNATQA"
BASE COUNT      205 a 230 c 262 g 233 t
ORIGIN
```

```
Query Match      6.4% Score 76.6; DB 6; Length 930;
Best Local Similarity 54.7%; Pred. No. 1.5e-08;
Matches 179; Conservative 0; Mismatches 139; Indels 9; Gaps 1;
OY 399 TAGTGAATGTTGGGTGTGTCATAGTGTGCGACCTAGATACGTGCGCAAGAAATTTGGAG 458
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 248 TATTGATGATGCGTTGGCGCATTTGGCGGTAGACAGTCGATCTCGCCAAATATATGTT 307
OY 459 CAACCAAGTGAAGCATCTGATCTGATCTGTTCAGACTCAAGAAGCAATATCTTGGCTG 518
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 308 TTAAGCTTACCGGCATCCATTAAGTCCCAACAGGTGAACAGGGCGACGAAATTAATCT 367
OY 519 CTGCTCAAGATGTGGCTATAGGTTCTTCAAGGTGTGACGCTTACAGCAACCAT 578
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 368 CTCCGATGACGCGCAAG-----TTTGGCGTGAGAGATGCTATGCTTGTCTT 418
OY 579 TCTCTAGCGCCAGTTTGATCTGTGTCATGAGAGAGTGAAGCATATGCTTGACA 638
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 419 TTCTGAGCGTAGTTTGCACATAGTTGTGTGGTGAAGAGAGGCGCCACATGCTGACA 478
OY 639 AAGCTAAGTTTGTGGAGAGTATGCTGCGGTAGACAGCACAGGTCCTTATATATAG 698
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 479 AAGCTGTGTTGCCAAGAAATTAAGTCCGCGGTGTAACACAGGGGCGATCTGTGGTGG 538
OY 699 TAACATGTGGCCACAGAGATCTTGCC 725
DB 539 CGGATTGGAATCAAGGAGCATGCGC 565
```

```
RESULT 13
AX073657      957 bp      DNA      linear      PAT 06-FEB-2001
LOCUS
DEFINITION      Sequence 1 from Patent W00104330.
ACCESSION      AX073657
VERSION      AX073657.1 GI:12710078
KEYWORDS
SOURCE      Synecocystis sp. PCC 6803.
ORGANISM      Synecocystis sp. PCC 6803
REFERENCE      Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
AUTHORS      1 (bases 1 to 957)
TITLE      Herbers, K., Badur, R., Kunze, I. and Geiger, M.
              Identification and overexpression of a dna sequence coding for
              2-methyl-6-phytylhydroquinone-methyltransferase in plants
```

Query Match	6.4%	Score 76.6	DB 6	Length 957
Matches 179	Conservative 0	Mismatches 139	Indels 9	Gaps 1
QY 399	TAGTGTGATGTTGGTGTGTGTCATAGCGTGCAGCTCTAGATACCTGTGCCAAGAATTTGGAG	458		
Db 290	TATTGGATGTGGGCTTGGCGCATTTGGCGGATGACAGTTCGGATTCTGCCAAGAATTATGCTT	349		
QY 459	CAACCATGTGATGCGCATCTCTGTAATCTCTGTTCAAAGCTTCAAAGACAAATGCTCTTGGTG	518		
Db 350	TTAAAGCTTACCGCGCATCACATTACTGCCCAACAGTGAAGGGCGCAGCAATTAATCTC	409		
QY 519	CTGCTCAGGATTTGGCTGTATAGGTTTCTTTTCAAGTTGGCTGACGCTCTACAGCAACAT	578		
Db 410	CTCCCGATGTGACGGCCAG-----TTTGGCGTGTGACGATCTTATGCTTGTCTT	460		
QY 579	TCTGTGACGGCCAGTTTGTGTGTGTGTCTCCATGAGAGATGAGAGCATATGCTCTGACA	638		
Db 461	TTCTCTGACGGCTATGTTTCCACGATGTTTGGTGTGGATGGAACAGAGGCCCCACATGCTTGACA	520		
QY 639	AAGCTAAGTTTGTGGAGATTAGCTCGGGTAGACAGCACCAGGTGCCATTATTAATAATAG	698		
Db 521	AAGCTGTGTTTCCCAAGCAATTAATCTGCGGGTGTGTAAACACAGGGGCGATTCTGTGTGTGG	580		
QY 699	TAACTGTGTGCCACAGGATCTTGGCC	725		
Db 581	CGGATTGGAATCAACGGCAGATCGCC	607		

BASE COUNT	203 a	244 c	268 g	242 t
ORIGIN				

transl_table=1	/protein_id="CA035426.1"
/db_xref="GI:2152287"	
/translation="MPEYLLPAGLISLSLAIAAGLYLTARGTSSDSVANAYDQMT	
EDGLEIETWGDHILHGHYGDPPVAKDIQSITDVFHAMAGGIDTLPFGTIVLDVGC	
GIGSSRIILAKDYENFTGITISPOQVRAKTELPDYTAKEFAVDAMALSPGSGSD	
VVMSVEAGPHMDKAVFAKELLRVKPGGLLVADMDRDRQVPLNFMKRPVROLLL	
DQMSHPAFASIEGFAENLEATGIEGVGTATMDVPTLPAIMDLTWGIIIRPGQWLYQ	
GIRGFIKSVREVPITLLMRILAFVGLCRFGRFKVRKNATQA"	

LOCUS	AX073659	974 bp	DNA	linear	PLAT 06-FEB-2001
DEFINITION	Sequence 3 from Patent WO0104330.				
ACCESSION	AX073659				
VERSION	AX073659.1				
KEYWORDS	GI:12710080				
SOURCE					
ORGANISM	Synechocystis sp. PCC 6803.				
REFERENCE	Synechocystis sp. PCC 6803				
AUTHORS	Bacteria; Cyanobacteria; Chroococcales; Synechocystis.				
TITLE	1 (bases 1 to 974)				
JOURNAL	Hebers,K., Badur,R., Kunze,I. and Geiger,M.				
FEATURES	Identification and overexpression of a dna sequence coding for				
SOURCE	2-methyl-6-phytylhydroquinone-methyltransferase in plants				
	Patent: WO 0104330-A 3 18-JAN-2001;				
	Location/Qualifiers				
	1..974				
	/organism="Synechocystis sp. PCC 6803"				
	/db_xref="taxon:1148"				
	7..963				
	/note="unnamed protein product"				
	/codon_start=1				
	/transl_table=11				
	/protein_id="CAC28416.1"				
	/db_xref="GI:12710081"				
	/translation="MPEYLLPAGLISLSLAIAAGLYLTARGTSSDSVANAYDQMT				
	EDGLEIETWGDHILHGHYGDPPVAKDIQSITDVFHAMAGGIDTLPFGTIVLDVGC				
	GIGSSRIILAKDYENFTGITISPOQVRAKTELPDYTAKEFAVDAMALSPGSGSD				
	VVMSVEAGPHMDKAVFAKELLRVKPGGLLVADMDRDRQVPLNFMKRPVROLLL				
	DQMSHPAFASIEGFAENLEATGIEGVGTATMDVPTLPAIMDLTWGIIIRPGQWLYQ				
	GIRGFIKSVREVPITLLMRILAFVGLCRFGRFKVRKNATQA"				

BASE COUNT 206 a 249 c 273 g 246 t
ORIGIN

Query Match 6.4%: Score 76.6; DB 6; Length 974;

Best Local Similarity 54.7%: Pred. No. 1.5e-08;
Matches 179; Conservative 0; Mismatches 139; Indels 9; Gaps 1;

```
OY 399 TAGTGATGTGGGTGGCATAGTGGCAGCTAGATACCTGGCCAGAGAATTGGAG 458
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 296 TATTGATGTGGGTGGCATAGTGGCAGCTAGATACCTGGCCAGAGAATTGGAG 355
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 459 CAACCAAGTGAAGCATCACTGAGTCCCTGTTCAAGCTCAAGAGCAATGCTTGTG 518
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 356 TTAACGTTACCGGCAATCACTGAGTCCCTGTTCAAGCTCAAGAGCAATGCTTGTG 415
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 519 CTGCTCAAGATGGCTGATTAAGTTCCTTTCAGAGTGTGAGCGCTTACAGCAACAT 578
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 416 CTCCGATGTAGCGGCAAG-----TTGCGGTGAGCGATGCTATGGCTTGTCTT 466
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 579 TCTGTAGCGGCAATGATCTGTGTGTCATGTGAGAGTGAAGCATATGCTGACA 638
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 467 TTCTGACGAGTATGTTGACGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 526
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 639 AAGCTAAGTGTGTTGAGAGTTAGCTCGGAGTGAAGCAACAGGATGATTAATATAG 698
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 527 AAGCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 586
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 699 TAACTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 725
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 587 CGGATTTGATCAACGAGGATGATGATGATGATGATGATGATGATGATGATGATG 613
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: March 18, 2003, 08:23:32
Job time : 3728 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2003, 06:42:05 ; Search time 1529 Seconds

(without alignments)
1751.227 Million cell updates/sec

Title: US-09-857-613A-27

perfect score: 1189
Sequence: 1 ggcacatggccaccgtgtgta.....aaaaaaaaaaaaaaaa 1189

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : N.Geneseq.101002:*

- 1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1189	100.0	1189	21	AAA29164
2	953	80.2	1331	21	AAA29154
3	459	38.6	1257	21	AAA29165
4	445.2	37.4	1293	21	AA050535
5	442	37.2	1071	21	AA298317
6	442	37.2	1350	24	AA146039
7	442	37.2	1790	20	AA117789
8	440.4	37.0	1296	21	AA034022
9	436.2	36.7	1102	21	AAA29161

10	353.2	29.7	1011	21	AAA29155	Wheat gamma tocoph
11	322.2	27.1	1862	22	AAH44261	Physcomitrella pat
12	310	26.1	792	21	AAA29151	Corn gamma tocophe
13	147	12.4	954	20	AA017788	Synechocystis gamm
14	125.8	10.6	521	21	AAA29152	Partial gamma toco
15	125.8	10.6	521	21	AAA29162	Rice gamma-tocophe
16	124.8	10.5	311	24	ABL74098	Corn tassal-deri
17	95.8	8.1	464	21	AAA29163	Rice gamma-tocophe
18	79.6	6.7	488	22	AAH44245	Physcomitrella pat
19	76.6	6.4	930	22	AA026173	Synechocystis PCC6
20	76.6	6.4	957	21	AA261599	DNA encoding a met
21	76.6	6.4	957	22	AA026169	Synechocystis PCC6
22	76.6	6.4	957	24	AA046040	Synechocystis 2-me
23	76.6	6.4	974	22	AA026170	Synechocystis PCC6
24	72.8	6.1	592	21	AAA29153	Rice gamma tocophe
25	50	4.2	60	21	AAA29170	Forward primer to
26	47.8	4.0	1410	15	AA080524	Oxidoreducing aver
27	47.8	4.0	7143	24	ABL32983	Human immune syste
28	46.8	3.9	12381	21	AA258381	Streptomyces averm
29	45.4	3.8	752	21	AA068123	Human secreted pro
30	45.4	3.8	1923	21	AAA93825	Modified fibre pro
31	45.4	3.8	6922	22	AA046533	Chemically treated
32	45.4	3.8	6922	24	AA061331	Human gene regulat
33	45.4	3.8	6922	24	ABK31410	Signal transductio
34	45.4	3.8	6922	24	AAH34576	Human colon cancer
35	45.2	3.8	521	22	AA055802	S. lavendulae Mltm
36	44.2	3.7	852	21	AA055842	Complete nucleotid
37	44.2	3.7	53500	21	AA055842	Human prostate exp
38	43.8	3.7	277	23	ABV18238	Human prostate exp
39	43.6	3.7	2045	21	AA016314	Human prostate can
40	43.2	3.6	405	22	AA184621	Human polynucleoti
41	43.2	3.6	466	23	ABV05510	Human prostate exp
42	43.2	3.6	475	23	ABV48025	Human prostate exp
43	43.2	3.6	556	23	ABV40063	Human prostate exp
44	43.2	3.6	556	23	ABV40163	Human prostate exp
45	43.2	3.6	556	23	ABV42105	Human prostate exp

ALIGNMENTS

RESULT 1	AAA29164	standard; cDNA: 1189 BP.
ID	AAA29164	
AC	AAA29164:	
XX		
DT	12-SEP-2000	(first entry)
XX		
DE	Soybean gamma-tocopherol methyltransferase cDNA (clone sahic.pk004.92).	
XX		
KW	Vitamin E: alpha-tocopherol: biosynthesis; enzyme; inhibitor; herbicide;	
KW	gamma-tocopherol methyltransferase; ss.	
XX		
OS	Glycine max.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	5..1057
FT		/tag= a
FT		/product= gamma-tocopherol_methyltransferase
XX		
PN	MO200032757-A2.	
PD	08-JUN-2000.	
XX		
XX	02-DEC-1999;	99WO-US28588.
XX		
PR	03-DEC-1998;	98US-0110781.
XX		
PA	(DUPO) DU PONT DE NEMOURS & CO E. I.	
XX		
XX	Arabidopsis gamma-t	
XX	Arabidopsis gamma-t	
XX	Arabidopsis thaila	
XX	Cahoon RE, Coughlan SJ, Miao G, Rafalski JA;	
XX		

xx Claim 3: Page 52: 82pp: English.

cc AAA29151-69 encode vitamin E (alpha-tocopherol) biosynthetic enzymes or
cc their fragments. The cDNA clones were identified by BLAST searches based
cc on similarity to *Synechocystis* sp. and *Arabidopsis thaliana* vitamin E
cc biosynthetic enzymes. The enzymes are useful for synthesizing
cc plastoquinones or tocopherols (especially vitamin E). Vitamin E is
cc required for plant growth, therefore the enzymes may be used for the
cc discovery of new herbicides. The enzymes can be used in methods to
cc evaluate potential inhibitors, which may have use as herbicides.
cc Additionally, transgenic expression of, e.g., gamma-tocopherol
cc methyltransferase affords the ability to manipulate tocopherol levels as
cc desired for a particular application.

xx Sequence 1331 BP; 373 A; 293 C; 301 G; 364 T; 0 other;

Query Match 80.2%; Score 953; DB 21; Length 1331;

Best Local Similarity 100.0%; Pred. No. 1.5e-250;
Matches 953; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 2 GACATGGCCACCGTGTAGATATCCCAACATCTCATGATCCACATCCACATCCCT 61
Db 3 GACATGGCCACCGTGTAGATATCCCAACATCTCATGATCCACATCCACATCCCT 62
QY 62 TCCCAATCCCTCGCAGCTTTGCGCAGAAATCCGGTCGACCCAGGTCGTGGCTCTATT 121
Db 63 TCCCAATCCCTCGCAGCTTTGCGCAGAAATCCGGTCGACCCAGGTCGTGGCTCTATT 122
QY 122 CGGCGATCGGCGACGAGCTCGAGAGAGGGGAGATAGTATTTGAGCAGAGCCGAGAG 181
Db 123 CGGCGATCGGCGACGAGCTCGAGAGAGGGGAGATAGTATTTGAGCAGAGCCGAGAG 182
QY 182 GATGCAAGAGAACTGCGAAGGGAATCGCAGATTTTACGACAGTCGTCTGCTTA 241
Db 183 GATGCAAGAGAACTGCGAAGGGAATCGCAGATTTTACGACAGTCGTCTGCTTA 242
QY 242 TGGGAGAACTTTGGGGCGACACATGACACATGCTTTATGACTCGGATTCACATGTT 301
Db 243 TGGGAGAACTTTGGGGCGACACATGACACATGCTTTATGACTCGGATTCACATGTT 302
QY 302 TCGCTTTGGATCATGCTGCTCTCAGATCCGAATGATCCAGAGATCTCTCGCTTGGC 361
Db 303 TCGCTTTGGATCATGCTGCTCTCAGATCCGAATGATCCAGAGATCTCTCGCTTGGC 362
QY 362 TCTGTTTCTGAGAGCGTGTAAATGCGCCAGAGATATGATTTGGGTGGCATA 421
Db 363 TCTGTTTCTGAGAGCGTGTAAATGCGCCAGAGATATGATTTGGGTGGCATA 422
QY 422 GGTGGCAGCTAGATACCTGGCCAAAGAAATTTGAGCAACGATGATGATCCTCTG 481
Db 423 GGTGGCAGCTAGATACCTGGCCAAAGAAATTTGAGCAACGATGATGATCCTCTG 482
QY 482 AGTCTGTTCAAGCTCAAGAGCAAAATGCTTGTGCTGCTCAAGAGATTTGCTGATAG 541
Db 483 AGTCTGTTCAAGCTCAAGAGCAAAATGCTTGTGCTGCTCAAGAGATTTGCTGATAG 542
QY 542 GTTTCCTTTACAGTGTCTCAAGCTCTACAGCAACATTTCTCTAGCGCCAGTTGATCTG 601
Db 543 GTTTCCTTTACAGTGTCTCAAGCTCTACAGCAACATTTCTCTAGCGCCAGTTGATCTG 602
QY 602 GTTGGTCCATGAGAGGTGAGAGCATATGCCGACAAAGCTTAAGTTTGTGGAGGTTA 661
Db 603 GTTGGTCCATGAGAGGTGAGAGCATATGCCGACAAAGCTTAAGTTTGTGGAGGTTA 662
QY 662 GCTCGGTAAGAGACAGAGTGCATTATATATAGTAACATGTCGCCACAGGATCTT 721
Db 663 GCTCGGTAAGAGACAGAGTGCATTATATATAGTAACATGTCGCCACAGGATCTT 722
QY 722 GGCCTGACGAACATCTTACATCCATGGGAGCAAGATCTTTAAAGAAATTTGGCAT 781
Db 723 GGCCTGACGAACATCTTACATCCATGGGAGCAAGATCTTTAAAGAAATTTGGCAT 782

```

```

QY 782 GCATATTTACCTCCCTGCTGCTGCTCAACTCTGTATATGTTAAGTGTCTCCATCCCTG 841
Db 783 GCATATTTACCTCCCTGCTGCTGCTCAACTCTGTATATGTTAAGTGTCTCCATCCCTG 842
QY 842 TCACTTCAGACATCAAGTCAGACAAATTTGCTGCTTGTGCTCCATTGGCCAGCA 901
Db 843 TCACTTCAGACATCAAGTCAGACAAATTTGCTGCTTGTGCTCCATTGGCCAGCA 902
QY 902 GTGATACGCTCAAGCTTCACATGAGAAAGGCTATCTTCACTGAGCAGTGG 954
Db 903 GTGATACGCTCAAGCTTCACATGAGAAAGGCTATCTTCACTGAGCAGTGG 955

```

RESULT 3
AAA29165
ID AAA29165 standard; cDNA; 1257 BP.

AC AAA29165;
XX
XX 12-SEP-2000 (first entry)
XX

DE Soybean gamma-tocopherol methyltransferase cDNA.

KW Vitamin E: alpha-tocopherol; biosynthesis; enzyme; inhibitor; herbicide;
KW gamma-tocopherol methyltransferase; ss.

XX Triticum aestivum.

XX Key Location/Qualifiers
XX CDS 18..1118

FT /tag= a
FT /transl_except= (pos:30..32, aa:xaa)
FT /transl_except= (pos:150..152, aa:xaa)
FT /product= gamma-tocopherol methyltransferase
FT /note= "xaa is not defined"

PN WO200032757-A2.

PD 08-JUN-2000.

PF 02-DEC-1999; 99WO-US26588.

XX 03-DEC-1998; 98US-0110781.

PA (DUPO) DU PONT DE NEMOURS & CO E I.

PI Cahoon RE, Coughlan SJ, Miao G, Rafalski JA;

DR WPI: 2000-412309/35.

XX P-PSDB; AAY96473.

PT polynucleotide encoding a vitamin E biosynthetic enzyme polypeptide
PT such as gamma tocopherol methyltransferase, useful for synthesizing
PT vitamin E or as a target for design and discovery of herbicides

XX Claim 3: Page 66: 82pp: English.

cc AAA29151-69 encode vitamin E (alpha-tocopherol) biosynthetic enzymes or
cc their fragments. The cDNA clones were identified by BLAST searches based
cc on similarity to *Synechocystis* sp. and *Arabidopsis thaliana* vitamin E
cc biosynthetic enzymes. The enzymes are useful for synthesizing
cc plastoquinones or tocopherols (especially vitamin E). Vitamin E is
cc required for plant growth, therefore the enzymes may be used for the
cc discovery of new herbicides. The enzymes can be used in methods to
cc evaluate potential inhibitors, which may have use as herbicides.
cc Additionally, transgenic expression of, e.g., gamma-tocopherol
cc methyltransferase affords the ability to manipulate tocopherol levels as
cc desired for a particular application.

SO Sequence 1257 BP; 299 A; 370 C; 343 G; 243 T; 2 other;

Query Match 38.6%; Score 459; DB 21; Length 1257;
Best Local Similarity 70.9%; Pred. No. 2.6e-115;

Query Match	37.4%: Score 445.2; DB 21; Length 1293;
Best Local Similarity 70.5%; Pred. No. 1.6e-111;	
Matches 626; Conservative 0; Mismatches 253; Indels 9; Gaps 2	
PR 01-SEP-1999; 99US-0151930.	
PR 07-SEP-1999; 99US-0152363.	
PR 10-SEP-1999; 99US-0153070.	
PR 13-SEP-1999; 99US-0153758.	
PR 15-SEP-1999; 99US-0154018.	
PR 16-SEP-1999; 99US-0154039.	
PR 20-SEP-1999; 99US-0154779.	
PR 22-SEP-1999; 99US-0155139.	
PR 23-SEP-1999; 99US-0155486.	
PR 24-SEP-1999; 99US-0155659.	
PR 28-SEP-1999; 99US-0156458.	
PR 29-SEP-1999; 99US-0156596.	
PR 04-OCT-1999; 99US-0157713.	
PR 05-OCT-1999; 99US-0157753.	
PR 06-OCT-1999; 99US-0157865.	
PR 07-OCT-1999; 99US-0158028.	
PR 08-OCT-1999; 99US-0158232.	
PR 12-OCT-1999; 99US-0158369.	
PR 13-OCT-1999; 99US-0159293.	
PR 13-OCT-1999; 99US-0159294.	
PR 13-OCT-1999; 99US-0159295.	
PR 14-OCT-1999; 99US-0159329.	
PR 14-OCT-1999; 99US-0159330.	
PR 14-OCT-1999; 99US-0159331.	
PR 14-OCT-1999; 99US-0159637.	
PR 14-OCT-1999; 99US-0159638.	
PR 18-OCT-1999; 99US-0159584.	
PR 21-OCT-1999; 99US-0160741.	
PR 21-OCT-1999; 99US-0160767.	
PR 21-OCT-1999; 99US-0160768.	
PR 21-OCT-1999; 99US-0160770.	
PR 21-OCT-1999; 99US-0160814.	
PR 21-OCT-1999; 99US-0160815.	
PR 22-OCT-1999; 99US-0160980.	
PR 22-OCT-1999; 99US-0160981.	
PR 22-OCT-1999; 99US-0160989.	
PR 25-OCT-1999; 99US-0161404.	
PR 25-OCT-1999; 99US-0161405.	
PR 25-OCT-1999; 99US-0161406.	
PR 26-OCT-1999; 99US-0161359.	
PR 26-OCT-1999; 99US-0161360.	
PR 26-OCT-1999; 99US-0161361.	
PR 28-OCT-1999; 99US-0161920.	
PR 28-OCT-1999; 99US-0161992.	
PR 28-OCT-1999; 99US-0161993.	
PR 29-OCT-1999; 99US-0162142.	
Query Match	37.4%: Score 445.2; DB 21; Length 1293;
Best Local Similarity 70.5%; Pred. No. 1.6e-111;	
Matches 626; Conservative 0; Mismatches 253; Indels 9; Gaps 2	
OY 187 CAGAAGAGCTGCACAGAGGAATCCAGATTTCAGACGAGTCTTGCTTATGGCA 246	
DB 233 CACTGAGGGGCTAAGAAAGAAATAGCGGTTCTACAATGAAACTTCGGGTTGTGGCA 292	
OY 247 GAACATTTGGGGCGACCAATGCACCATGGCTTTATAGACTCGATTCACATGTTCCG 306	
DB 293 AGAGATTGGGGGATCATATGCATATGGCTTTATGACCCGTGATTCCTTTCAACT 352	
OY 307 TTGGGATCATCCG-----GCTGCTAGATCCGATGATCCAGAGTCTCTCTGCTTGC 360	
DB 353 TTCTGATTTGGGTGCACAGAGCTCAGATCCGATTCGATTTGAAGAGTCTCTCGTTTGC 412	
OY 361 CTCCTGTTTC---TGAGAGAGCTAGTAAATGGCCCAAGAGTATAGTTGATGTTGGTGTGG 417	
DB 413 CGGTGTACTGATGAGAGAGAGAGAGAAAGATAAAGATAGTAGTGGATTTGGGTGG 472	
OY 418 CATAGTGGCAGCTCTAGATACCTGGCCCAAGAAATTTGGAGCAACACAGTGTAGGCATCAC 477	
DB 473 GATTGGAGGAGAACTCAAGATATCTTGCTTAAATTTGGAGCGAATGCAATTTGGCAATTAC 532	
OY 478 TCTGAGATCCCTTCAAGCTCAAGAGCAAAATGCTCTGTGCTGCTCAAGATTTGGCTGA 537	

Db 956 GAAAGATTAAGAGCATGACAAATCCATGATGATTGAAGTTACAGAAAGGTGT 1015
 Oy 1018 AATTAGTTGGCATTCATTCATGTCGAAAGACCGAATTAATGAGAG 1065
 Db 1016 CATTAGTTGGTATCATCTGCGCAAGCCACTTAAGTCTTAAG 1063

RESULT 6
 AAL46039
 ID AAL46039 standard; cDNA: 1350 BP.
 AC AAL46039;
 XX 11-JUL-2002 (first entry)
 DE A thaliana gamma-tocopherol methyltransferase cDNA.
 XX Vitamin E; homogenisate; HG; homogenisate-1,2-dioxygenase; HGD;
 KW maleylacetate-isomerase; MAAI; fumarylacetoacetate-hydrolase; FAAH;
 KW transgenic; nutrition; antioxidant; cardiovascular disease; cancer;
 KW immune system; generalised age-related degeneration; animal feed;
 KW meat quality; cosmetics; growth regulator; herbicide; cardiant;
 KW cytosolic; immunostimulant; enzyme; gene; ss.
 XX Arabidopsis thaliana.
 OS
 FH Key Location/Qualifiers
 FT 63..1109
 FT CDS /tag= a
 FT /product= "gamma-tocopherol methyltransferase"
 XX
 PN MO200231173-A2.
 PD 18-APR-2002.
 PE 18-SEP-2001; 2001MO-EPI0779.
 XX
 PR 19-SEP-2000; 2000DE-1046462.
 PA (SUNG-) SUNGENE GMBH & CO KGAA.
 PI Gelger M, Ebneth M, Kunze I:
 DR WPI: 2002-362682/39.
 DR P-PSDB: AAO17426.
 XX
 PT Increasing Vitamin E production, useful in human or animal nutrition,
 PS comprises reducing activity of enzymes that metabolize homogenisate
 XX Disclosure: Page 83-85; 98pp; German.
 XX
 CC The present invention relates to a method of producing Vitamin E
 CC involving modulating the synthesis of vitamin E by reducing decomposition
 CC of homogenisate (HG) by reducing the activity of
 CC homogenisate-1,2-dioxygenase (HGD), maleylacetate-isomerase (MAAI)
 CC and/or fumarylacetoacetate-hydrolase (FAAH). The method is used to
 CC produce transgenic organisms that are useful in animal and human
 CC nutrition and for the isolation of vitamin E. Vitamin E is a fat-soluble
 CC antioxidant with a protective effect against cardiovascular disease and
 CC cancer. It also stimulates the immune system and may prevent generalised
 CC age-related degeneration. When used in animal feeds, it improves quality
 CC and storability of meat and can also be used in cosmetics. Also
 CC antibodies raised against the 3 specified enzymes are useful in screening
 CC for specific inhibitors, potentially useful as growth regulators, e.g.
 CC herbicides. The present sequence is the A thaliana gamma-tocopherol
 CC methyltransferase cDNA.
 XX
 SQ Sequence 1350 BP; 374 A; 278 C; 314 G; 384 T; 0 other;

Query Match 37.2%; Score 442; DB 24; Length 1350;
 Best Local Similarity 70.3%; Pred. No. 1.2e-110;
 Matches 624; Conservative 0; Mismatches 255; Indels 9; Gaps 2;

Oy 187 CAGAGAGCTGACAGAGGAATCCGAGAGTTTACAGAGCTGCTGCTATAGGA 246
 Db 230 CACTAGAGCGCTTAAGAAAGGAATAGCGAGTTTACAAATGAACCTCGGGTTGTGGGA 289
 Oy 247 GAACATTTGGGGCGACACATGACACATGCTTTTATGACTCGGATTCACCTGTTTCGT 306
 Db 290 AGAGATTTGGGAGATCATATGATTCATGATGGCTTTTATGACCCGTGATCTCTCTCAACT 349
 Oy 307 TTGCGATCATGCT-----GCTGCTAGATCCGATGATCCAGAGCTCTGCTTC 360
 Db 350 TTCTGATTTCTGCTCACAAGAGACCTCAGATCCGATGATGATGAAAGTCTCTCCCTTCGC 409
 Oy 361 CTCTGTTTC---TAGAGAGCTACTAATATGGCCCAAGATATGATGATGTTGGGTGGG 417
 Db 410 CGGTGTACTGATGATGAAGGAGGAGAAAGATTAAGAAAGTATGATGGTGGGTGGG 469
 Oy 418 CATAGGTGGCAGCTCTATACCTGAGCCCAAGAAATTTGGACCAACAGTATGAGCATCAC 477
 Db 470 GATTGGAGAGAGCTCAAGATATCTTGGCTCTAATTTGAGCTGAATGCAATGGCATTTG 529
 Oy 478 TCTGAGTCTGTTCAAGCTCAAGAGCAAAATGCTTCTGCTGCTCAAGATTTGGCTGA 537
 Db 530 TCTCAGCCCTGTTCAGGCCAAGAGCAATGATCTCGCGCTGCTCAATCACTCTCA 589
 Oy 538 TAAGGTTCTTTACAGTTGCTGACGCTCTACAGCAACCATTTCTGACGGCCAGTTTGA 597
 Db 590 TAAGGCTCTTCCCAAGTTGGGATGCGTTGATGATGAGCTGAAGATGAAGAAATTTGA 649
 Oy 598 TCTGCTGCTGCTCAGAGAGTGGAGCAATATGCTGCTCAAGAGCTAATTTGTTGAGA 657
 Db 650 TCTAGTGTGCTGATGAGAGAGTGTGAGCATATGCTGACCAAGGCCAAGTTTGAAGA 709
 Oy 658 GTTAGCTCGGAGAGCAGCAGTGCATTTAATTAATAGTAAATGATGATGATGATGATG 717
 Db 710 GTTGTAGCTGTGGGGCTCCAGAGATAGATTAATTAATGATGATGATGATGATGATG 769
 Oy 718 TCTTGCCCTGAGCAATCTTATCATCATGAGAGCAAGATCTCTTAAGAAATTTG 777
 Db 770 TCTATCTGCGGGGAGAGAGCTTTCAGCCGCGGAGCAAAATCTTGGACAAAATCTG 829
 Oy 778 CGATGATATATCTCCCTGCTGCTGCTCACTCTGATTAATGATGATGATGATGATG 837
 Db 830 TAAGAGTCTCTATCTCCGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 889
 Oy 838 CCTGCTCACTTCAAGCATCAAGTCAAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 897
 Db 890 CCAATCTTCCAGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 949
 Oy 898 AGCAGTATACGCTCAGCTTCAATGATGAAGGCTATCTTCACTCTTGAAGCATGAGCA 957
 Db 950 TGGGTTATACGAGCTGATTAATGATGAAGGCTTGTCTCTGCTGCTGCTGCTGCTGAT 1009
 Oy 958 AAAACGATTAAGAGAGCTTGGCTATGCTGATGATGATGATGATGATGATGATGATG 1017
 Db 1010 GAAAGATTAATTAAGAGAGCTTGAATGATGATGATGATGATGATGATGATGATGATG 1069
 Oy 1018 AATTAGTTGGCATTCATTCATGTCGAAAGACCGAATTAATGAGAG 1065
 Db 1070 CATTAGTTGGTATCATCTGCGCAAGCCACTTAAGTCTTAAG 1117

RESULT 7
 AAX17789
 ID AAX17789 standard; DNA: 1790 BP.
 AC AAX17789;
 XX 21-MAY-1999 (first entry)
 DE Arabidopsis gamma-tocopherol methyltransferase encoding gene.
 XX
 XX Gamma-tocopherol methyltransferase: gamma-TMT; S1R0089; 165H577; meat;
 KW transgenic plant; alpha-tocopherol; gamma-tocopherol; forage plant; ds.

```

XX OS Arabidopsis sp.
XX FH Key Location/Qualifiers
XX CDS 207..1253
XX FT /tag=
XX FT /gene="165H5T7"
XX FT /product="gamma-TMT"
XX PN
XX MO9904622-A1.
XX PD 04-FEB-1999.
XX PF 22-JUL-1998; 98MO-US15137.
XX PR 17-JUL-1998; 98US-0053819.
XX PR 25-JUL-1997; 97US-0053819.
XX PR 26-JAN-1998; 98US-0072497.
XX PA (UYNE-) UNIV NEVADA.
XX PI DeJapenna D, Shintani DK;
XX DR MPI: 1999-142458/12.
XX DR P-PSDB; AAM95017.
XX PT Newly isolated DNA fragment comprising a ^c-tocopherol (vitamin E)
XX PT methyltransferase coding sequence - useful for producing
XX PT ^a-tocopherol, and transgenic plants, seeds and oils with an altered
XX PT tocopherol profile
XX PS Claim 2; Page 36-38; 46pp; English.
XX XS
XX CC The invention provides DNA sequences encoding gamma-tocopherol
XX CC methyltransferases (gamma-TMT). The gamma-TMT encoding genes SHR0089
XX CC and 165H5T7 are isolated from Synchocystis and Arabidopsis species
XX CC respectively. The DNA fragments are useful for producing transgenic
XX CC plants containing a higher alpha-tocopherol:gamma-tocopherol ratio. They
XX CC are also useful for producing alpha-tocopherol and plants with increased
XX CC gamma-tocopherol, which may be useful in certain industries such as the
XX CC meat industry e.g. for developing forage plants to feed animals. The
XX CC production of transgenic plants (and seeds) with a higher alpha-
XX CC tocopherol:gamma-tocopherol ratio, using the gamma TMTs, will increase
XX CC the level of alpha-tocopherol in the human diet, in addition to enhancing
XX CC the stability and shelf life of plants and plant products. Increased
XX CC levels of alpha-tocopherol will also increase meat quality and extend
XX CC shelf life of post-processed meat products. Plants with a higher alpha-
XX CC tocopherol:gamma-tocopherol ratio may also produce advantageous
XX CC phenotypes. The present sequence represents the 165H5T7 gene encoding the
XX CC Arabidopsis gamma-TMT.
XX SQ Sequence 1790 BP; 474 A; 404 C; 436 G; 476 T; 0 other;
XX
XX Query Match 37.28; Score 442; DB 20; Length 1790;
XX Best Local Similarity 70.38; Pred. No. 1.4e-110;
XX Matches 624; Conservative 0; Mismatches 255; Indels 9; Gaps 2;
XX
OY 187 CAAGAGAGAGCTCCAGAGGATCCAGAGTTTACGACGAGTCTGCGCTTATGGA 246
DB 374 CACTGAGGGCGCTAAGAAAAGGATACGGAGTCTTACATGAAGAACTCGGTTG 433
OY 247 GAACATTTGGGGCGACACATGACCATGCTTTTATGACTGCGATTCACCTTTTGC 306
DB 434 AGAGATTGGGGAGATCATATGATGATGATGATGATGATGATGATGATGATGAT 493
OY 307 TTGCGATCATCTGCT-----GCTGCTCAGATCCGAATGATCAAGAGTCTTGCCTTGC 360
DB 494 TTCTGATTTCTGCTCAAGAGAGCTCAGATCCGATGATGATGATGATGATGATGAT 553
OY 361 CTCCTGTTTC---TGAGAGAGCTAGTAAGATGAGCCCAAGATATAGTATGATGATGATG 417
DB 554 CGGTCTTACTGATGAAGAGAGGAGAGAAAGATTAAGAAAGATGATGATGATGATGATG 613

```

```

OY 418 CATAGTGGGACGCTCTAGATACCTGCGCCAGAAATTTGGACCAACAGTGGACATCAC 477
DB 614 GATTGAGAGAGAGCTCAGATATCTTCCCTTAATTTGGAGCTGAATGATGATGATGAT 673
OY 478 TCTGATGCTGTTTCAAGCTCAAGAGCAAAATGCTTGGCTGCTCCTCAAGGATTTGGCTGA 537
DB 674 TCTGAGCCCTGTTTCAAGCTCAAGAGCAAAATGATCTGCGGCTGCTCAATGATCTCTCA 733
OY 538 TAAGTTTCTTTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 597
DB 734 TAAGCTTCTTTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 793
OY 598 TCTGATGCTGTTTCAAGCTCAAGAGCAAAATGCTTGGCTGCTCCTCAAGGATTTGGCTGA 657
DB 794 TCTGATGCTGTTTCAAGCTCAAGAGCAAAATGCTTGGCTGCTCCTCAAGGATTTGGCTGA 853
OY 658 GTTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 717
DB 854 GTTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 913
OY 718 TCTTGGCCCTGAGCAAAATCCTTACATCCATGAGAGCAAGATCTTAAAGATTTG 777
DB 914 TCTATCTGCGGGGAGAGCTTTGACCGCTGAGCAAAACATCTTGGACAAATCTG 973
OY 778 CGATGATATATACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 837
DB 974 TAAGAGCTTCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1033
OY 838 CCTGTCACCTTCAAGAGATCAAGTCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 897
DB 1034 CCATCTCTCCAGATATTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1093
OY 898 AGCAGTATAGCTCAGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 957
DB 1094 TCGGTTATAGGAGCTCATTAAACATGAAAGGCTTGTCTGCTGCTGCTGCTGCTGCTGCTG 1153
OY 958 AAAAAGATTAAGAGAGCTTTGCTGATGCTGATGATGATGATGATGATGATGATGATGATG 1017
DB 1154 GAAAGATTAAGAGAGCTTTGCTGATGCTGATGATGATGATGATGATGATGATGATGATG 1213
OY 1018 AATTAGTTTGCATCATTTACATGTCGAAAAACCTGATTAATGAGAG 1065
DB 1214 CATTAAGTTTGTATCATCATCTGCAAGAGCACTTAAGCTTAAG 1261
XX
XX RESULT 8
XX AAC34022
XX ID AAC34022 standard; DNA; 1296 BP.
XX AC AAC34022;
XX DT 17-OCT-2000 (first entry)
XX XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 5167.
XX KM Hybridisation assay; genetic mapping; gene expression control;
XX KM protein identification; signal transduction pathway;
XX KM metabolic pathway; promoter; termination sequence; ss.
XX OS Arabidopsis thaliana.
XX PN BP1033405-AZ.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0123788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.

```

PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140354.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0143390.
PR 08-JUL-1999; 99US-0143803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 13-JUL-1999; 99US-0143624.
PR 14-JUL-1999; 99US-0144005.
PR 15-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151308.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.

CC methyltransferase affords the ability to manipulate tocopherol levels as
 CC desired for a particular application.
 XX
 SO Sequence 1102 BP; 231 A; 330 C; 328 G; 207 T; 6 other;

Query Match 36.7%; Score 436.2; DB 21; Length 1102;
 Best Local Similarity 69.9%; Pred. No. 4.2e-109;
 Matches 604; Conservative 0; Mismatches 254; Indels 6; Gaps 1;

```

OY 197 CTGCAGAGGGAATGCGAGTTTACGACGAGTGTGCTGCTTAAAGGAGAACATTTGG 256
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 187 CTGAGAGGAGCGATCGCGGCTGACGACGAGTGTGCTGAGGAGACATTCG 246
OY 257 GCGCAGCAGATCGACCATGCTTTTATGACGCGATTCACCTGTTTGGTTCGATCAT 316
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 247 GCGGACCAATCCACACGCGCTTACGACGTCGAGGAGCGCGCTCCATGGCGATCAC 306
OY 317 CGTGTCTGCTAGATCCGATGATTCACAGAGTCTCTTGGCTTGGCTCTGT-----TTCT 370
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 307 CGCCCGCGCGATCCGATCCGATGATTCGAGGCGCGCTCGCTTCCGCTCCAGCCTCA 366
OY 371 GAGGAGCGTAAATGCGCCCAAGATATGATGTTGATGGTGGGATAGGTGGCAGC 430
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 367 GATGATCCAGAGAGACACCAAAACAATGATGATGATGATGATGATGATGATGATGATG 426
OY 431 TCTAGATACCTGCGCCAGAAATTTGAGCAGCAGTGTAGGATCATCTGCTCTGTT 490
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 427 TCAAGATGATCTGCGAGAAATATGCGAGANGCAGTGCATGCGATGATGATGATGATGATG 486
OY 491 CAAGCTCAAGAGCAAAATCTCTTGTGCTGCTCAAGATGATGATGATGATGATGATGATG 550
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 487 CAAGCGGAGAGAAATCTCTGCTGCTGAGCGCGATGATGATGATGATGATGATGATGATG 546
OY 551 CAGTGTGCTGAGGCTCTAAGACCATCTCTGAGCGGCGATGATGATGATGATGATGATGATG 610
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 547 CAGTGTGCTGATGCTGTGAGCACTCTTCTGAGCGGAGTGTGATGATGATGATGATGATG 606
OY 611 ATGAGAGAGTGAAGCATATGCTTGAACAAAGTAACTTTGTTGAGAGATGATGATGATGATG 670
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 607 ATGAGAGAGTGAAGCATATGCTTGAACAAAGTAACTTTGTTGAGAGATGATGATGATGATG 666
OY 671 GCAGACACAGTGCCTATTAATATATGATGATGATGATGATGATGATGATGATGATGATG 730
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 667 GCGGCTCTGAGGAGCAATATCATGATGATGATGATGATGATGATGATGATGATGATGATG 726
OY 731 GAACATCTCTTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 790
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 727 GAACCTCTCTTAAAGCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 786
OY 791 CTCCTGCTGCTGCTCAACTTCTGATTAATGATGATGATGATGATGATGATGATGATGATGATG 850
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 787 CTCCTGCTGCTGCTCAACTTCTGATTAATGATGATGATGATGATGATGATGATGATGATGATG 846
OY 851 GACATCAATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 910
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 847 GATATCAACACAGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 906
OY 911 TCAGCTTCAATGAGAGGCTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 970
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 907 TCAGCTTCAATGAGAGGCTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 966
OY 971 GAGCTTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1030
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 967 GCGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1026
OY 1031 ATCATTACATGTCGAAAACCTGAA 1054
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1027 ATCATCACCCTGCGAAGCCTGGA 1050
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 10
 AAA29155
 ID AAA29155 standard; cDNA; 1011 BP.

```

XX AC AAA29155;
XX 12-SEP-2000 (first entry)
XX DE wheat gamma tocopherol methyltransferase cDNA.
XX KW Vitamin E; alpha-tocopherol; biosynthesis; enzyme; inhibitor; herbicide;
XX KW gamma-tocopherol methyltransferase; ss.
XX OS Triticum aestivum.
XX EH Key
XX EH CDS
XX FT Location/Qualifiers
XX FT 3..885
XX FT /tag=a
XX FT /partial
XX FT /product= gamma tocopherol methyltransferase
XX FT /transl_except= (pos:381..383, aa:Xaa)
XX FT /transl_except= (pos:387..389, aa:Xaa)
XX FT /transl_except= (pos:390..392, aa:Xaa)
XX FT /transl_except= (pos:393..395, aa:Xaa)
XX FT /transl_except= (pos:399..401, aa:Xaa)
XX FT /note="Xaa is not defined"
XX PN MO200032757-A2.
XX PD 08-JUN-2000.
XX PF 02-DEC-1999; 99WO-US28588.
XX PR 03-DEC-1998; 98US-0110781.
XX PA (DUP0) DU POINT DE NEMOURS & CO E. I.
XX PI Cahoon RE, Coughlan SJ, Miao G, Rafalski JA;
XX DR WPI; 2000-412309/35.
XX DR P-PSDB; AAY96463.
XX PT polynucleotide encoding a vitamin E biosynthetic enzyme polypeptide
XX PT such as gamma tocopherol methyltransferase, useful for synthesizing
XX PT vitamin E or as a target for design and discovery of herbicides
XX PS Claim 3; Page 53; 82pp: English.
XX CC AAA29151-69 encode vitamin E (alpha-tocopherol) biosynthetic enzymes or
XX CC their fragments. The cDNA clones were identified by BLAST searches based
XX CC on similarity to Synecchocystis sp. and Arabidopsis thaliana vitamin E
XX CC biosynthetic enzymes. The enzymes are useful for synthesizing
XX CC plastoquinones or tocopherols (especially vitamin E). Vitamin E is
XX CC required for plant growth, therefore the enzymes may be used for the
XX CC discovery of new herbicides. The enzymes can be used in methods to
XX CC evaluate potential inhibitors, which may have use as herbicides.
XX CC Additionally, transgenic expression of, e.g. gamma-tocopherol
XX CC methyltransferase affords the ability to manipulate tocopherol levels as
XX CC desired for a particular application.
XX SO Sequence 1011 BP; 255 A; 269 C; 268 G; 216 T; 3 other;

Query Match 29.7%; Score 353.2; DB 21; Length 1011;
Best Local Similarity 65.5%; Pred. No. 2.1e-86;
Matches 545; Conservative 0; Mismatches 285; Indels 2; Gaps 2;

```

```

OY 225 ACGAGTGTGCTGCTTATGAGAGCAATTTGGGCGAGCAGCATGACCATGCTTTATG 284
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 ACGAGTGTGCTGCTGCTTATGAGAGCAATTTGGGCGAGCAGCATGACCATGCTTTATG 60
OY 285 ACTGCGATTCATGCTTTGCTTGGATCATCTGCTGCTGATCCGATGATCCAG 344
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 61 ACTCCGGGAGGCGCTCCATGTCGACACCGCCCGCCAGATCGCATGATGATGAG 120
OY 345 AGTCTTCGCTTGGCTCTGTTCTGAGAGCGCTAGTAATGCGCCAGAGTATGTTG 404
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 121 AGGCCCCGCTTCGCCGCGCCGCCGAGATCCGACAAACCAACCAAGCATTTGTT 180
Oy 405 ATGTTGGGTGGGATAGTGGCAGCTCATAGATACCTGGCCAAAGAAATTTGGACACCA 464
Db 181 ATGTTGGATCGGAAATCGGTGGTAGCTCAAGATCTGGGCGCAAAATATGAGACAA 240
Oy 465 GGTAGGCACTGACTGACTGCTGCTCAAGTCAAGACCAATGCTTGGCTGCTGCTC 524
Db 241 TGTCTGGGATACATAGTACCCAGTGCAGAGTGAAGAGAAATGCTTCGGGACGCG 300
Oy 525 AAGATTTGGCTATAGTATTTCTTTCAGTTGCTGAGCTGACCTGACCAACCTTCCTG 584
Db 301 AAGGGGTTCGCGAACAAGTTCTTTCCATTGCTATCTGAGGACCAACCTTCCTG 360
Oy 585 ACGCCAGTTGATCTGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 643
Db 361 GATGGGATTTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Oy 644 AAGTTGTTGGAGATTTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 703
Db 421 AAGTTGTTAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Oy 704 TGTGTCACAGGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 763
Db 481 TGTGTCACATAGGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Oy 764 TTAAGAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 823
Db 541 TTAAGAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Oy 824 AAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 883
Db 601 AAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Oy 884 GCTTCATTTTGGCCAGCAGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 943
Db 661 GCTTCATTTTGGCCAGCAGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Oy 944 TTAGAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1003
Db 721 TTAAGAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Oy 1004 TTAAGAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1054
Db 781 TTAAGAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 832

RESULT 11
AAH44261
ID AAH44261 standard: DNA; 1862 BP.
AC AAH44261;
XX
XX 21-SEP-2001 (first entry)
DE Physcomitrella patens 78_ppp1.087_e12-259rev gene.
XX
XX Tocopherol and carotenoid metabolism related protein; TCMP; synthesis;
KW Physcomitrella patens; moss; algae; microorganism; fungus; plant;
KW identification; genome mapping; modulation; evolutionary study;
KW cellular production; fine chemical; ds.
XX
XX Physcomitrella patens.
XX
XX WO200144276-A2.
XX
XX 21-JUN-2001.
XX
XX 14-DEC-2000; 2000WO-EPI2698.
XX
XX 16-DEC-1999; 99US-0171121.
XX
XX (BADI) BASF PLANT SCI GMBH.
PA

XX
PI Lerchl J, Renz A, Ehrhardt T, Reindl A, Cirpus P, Bischoff F;
PI Frank M, Freund A, Duwenig E, Schmidt R, Reski R, Badur R;
DR WPI: 2001-398121/42.
DR P-PSDB; AAB99888.
XX
XX Tocopherol and carotenoid metabolism related protein (TCMP), used to
PT produce fine chemicals, is isolated from mosses, algae, microorganisms,
PT fungi, plants, or their fragments
XX
XX Claim 6; Page 111-112; 123pp; English.
XX
XX The present invention describes isolated tocopherol and carotenoid
CC metabolism related proteins (TCMP) (1) from mosses or algae,
CC microorganisms or fungi, plants, or its fragments. (1) can be used as
CC enzymes in the production of fine chemicals or in the metabolism of
CC tocopherols and carotenoids. (1) also assist in transmembrane transport.
CC The fine chemicals that can be produced include lipids, fatty acids,
CC vitamins, cofactors, enzymes, amino acids, and nucleotide bases.
CC Nucleotide sequences, proteins, vectors and host cells from the present
CC invention can be used: (a) to identify mosses related to Physcomitrella
CC patens; (b) in mapping genomes of mosses related to Physcomitrella
CC patens; (c) in the modulation of TCMP activity; (d) in evolutionary
CC studies; (e) in the determination of functional TCMP regions; (f) and
CC in the cellular production of fine chemicals. AAH44222 to AAH44262
CC encode the Physcomitrella patens TCMP proteins given in AAB99849 to
CC the exemplification of the present invention.
XX
XX Sequence 1862 BP; 465 A; 408 C; 521 G; 468 T; 0 other:
XX
XX Query Match 27.1%; Score 322.2; DB 22; Length 1862;
XX Best Local Similarity 61.4%; Pred. No. 8; 8e-78;
XX Matches 541; Conservative 0; Mismatches 328; Indels 12; Gaps 1;

Oy 189 AGAAGAGCTGCGAAGGAGATGCGAGATTTACGACAGTGTCTGCTTATGGGAGA 248
Db 377 AGCGGAGCTTTTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 436
Oy 249 ACATTTGGGGGCGACCATGACCATGCTTATGCTGCTGCTGCTGCTGCTGCTGCTG 296
Db 437 GCATTTGGGGGCGACCATGACCATGCTTATGCTGCTGCTGCTGCTGCTGCTGCTG 496
Oy 297 CTGTTGCGCTTTCGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 356
Db 497 TTGACGGGATCTGACACCGCGGAGCGCAATCAAGATGATGGAATCTCTGGCG 556
Oy 357 TTGCTCTGCTTTCGAGAGCGTATGTAATGCGCCCAAGATATGATGTTGGCTG 416
Db 557 ATGCTGGCGTTCCTGATAGCAAGATTTGAATGCAAGATGCTGCTGCTGCTGCTG 616
Oy 417 GCATAGTGGGAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 476
Db 617 GATAGGGGAGAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 676
Oy 477 CTCTGAGTCTGTTCAAGCTCAAGAGCAATGCTTCTGCTGCTGCTGCTGCTGCTG 536
Db 677 CGCTGACCCAGTGCAGAGGTTCAAGAGCGCTGACACCTTCTGCGCAAGAGCTTATG 736
Oy 537 ATAGGTTTCTTTCAGGTTGCTGAGCGCTTACAGACCATTTCTGAGCGCACTTG 596
Db 737 ACCTGCTCAATTTTCAGAGTACCAATGCTTGAACAGCGCTTTCAGAGTGTCTG 796
Oy 597 ATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 656
Db 797 ATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 856
Oy 657 AGTTAGCTGGGTAGAGAGCAAGTGTCCATTTTAATATAGTATGATGATGCTGCTG 716
Db 857 AGCTTGCAGAGTACAGAGCTCCCGGGGCTGCTATTCCTGCTGCTGCTGCTGCTG 916
Oy 717 ATCTTGGCCCTGACGAACAATCTTATCTCATGAGGAGCAAGATCTTAAAGAGATT 776

PN WO9094622-A1.
 XX
 PD 04-FEB-1999.
 XX
 PF 22-JUL-1998: 98WO-US15137.
 XX
 PR 17-JUL-1998: 98US-0053819.
 PR 25-JUL-1997: 97US-0053819.
 PR 26-JAN-1998: 98US-0072497.
 XX
 PA (UYNE-) UNIV NEVADA.
 XX
 PI Dellapenna D, Shintani DK;
 XX
 DR MPI: 1999-142458/12.
 XX P-PSDB: AAM95016.
 XX
 PT Newly isolated DNA fragment comprising a ^c-tocopherol (vitamin E)
 PT ^a-tocopherol, and transgenic plants, seeds and oils with an altered
 PT tocopherol profile
 PS
 PS Claim 2: Page 33-35; 46pp; English.
 XX
 CC The invention provides DNA sequences encoding gamma-tocopherol
 CC methyltransferases (gamma-TMT). The gamma-TMT encoding genes SLR0089
 CC and 165H57 are isolated from Synchocystis and Arabidopsis species
 CC respectively. The DNA fragments are useful for producing transgenic
 CC plants containing a higher alpha-tocopherol:gamma-tocopherol ratio. They
 CC are also useful for producing alpha-tocopherol and plants with increased
 CC gamma-tocopherol, which may be useful in certain industries such as the
 CC meat industry e.g. for developing forage plants to feed animals. The
 CC production of transgenic plants (and seeds) with a higher alpha-
 CC tocopherol:gamma-tocopherol ratio, using the gamma TMTs, will increase
 CC the level of alpha-tocopherol in the human diet, in addition to enhancing
 CC the stability and shelf life of plants and plant products. Increased
 CC levels of alpha-tocopherol will also increase meat quality and extend
 CC shelf life of post-processed meat products. Plants with a higher alpha-
 CC tocopherol:gamma-tocopherol ratio may also produce advantageous
 CC phenotypes. The present sequence represents the SLR0089 gene encoding a
 CC Synchocystis gamma-TMT.
 CC
 XX
 SO Sequence 954 BP; 204 A; 232 C; 271 G; 247 T; 0 other;
 Query Match 12.4%; Score 147; DB 20; Length 954;
 Best Local Similarity 49.9%; Pred. No. 6e-30; Mismatches 420; Indels 9; Gaps 2;
 Matches 428; Conservative 0;

Db 449 TGGCAATGCCCTTGGATTGGCCCTTCTCCGATTCCTTGGACTGGGTTGGTCGTGG 508
 Qy 615 AGAGTGGAGCATATGCTGACAAAGCTAAGTTGTTGGAGTATGCTGGGTACAG 674
 Db 509 AAGTGGGGAGACATGCCCAACAAAGCTATTTTACAGAAAGCTTGGGGGTACTTA 568
 Qy 675 CACCAAGTGCCTAATTAATAATAGTAACATGATGCGCACAGGATCTTGGCCCTACGAAC 734
 Db 569 AACCAAGTGCCTGCTGATTTTACGACCTGCTGCTACGTCATGCTCCATGATGCCGCAATG 628
 Qy 735 AATCCTACATCCATGGAGCAAGATCTTTAAAGAAATTTGGATGATATTACCTCC 794
 Db 629 GCCCCTGACTGCCGATGAACGTGCGCATCTCCAAAGCATCTATGAGCTTACTGTTGC 688
 Qy 795 CTGCTGTGCTCAACTTGTGATTAATGTTAAGTTGCTCCAAATCCCTCTACTCAGACA 854
 Db 689 CTTATGTTGTTTCCCTGCGGACTACGAGCGCATCCAGGAAATGTTGGGGGAAA 748
 Qy 855 TCAAGTCAGAAAGATTGCTGCTGCTTGTGCTCCATTTGGCCAGCAGTATAGCTCAG 914
 Db 749 TTAAGACTGCCGATGTTGTCAGTGGCGGTGACACTTTTGGACCGGCTGATGAGCTG 808
 Qy 915 CTTTACATGGAAGGCTATCTTCACTCTTGAACAGTGACAAAAACGATAAAGAG 974
 Db 809 CGTTGATCCCGGGGTTGTTGGGCTTTGGGCGCAAGCGGGCCAAAAATATCAATGCCG 868
 Qy 975 CTTTGGCTATGCTATGATGATGAGGATGATCAAGAAAGCTTAATTAATTTGCCATCA 1034
 Db 869 CCTGTGTTTACATTAATCAAAATGGGCTATGACGGGATTAATGCTGTTGGCTTAT 928
 Qy 1035 TTACATGTCGAAAACCT 1051
 Db 929 TTAAGGGGATTAAGCCT 945
 RESULT 14
 AAA29152
 ID AAA29152 standard; CDNA: 521 BP.
 AC AAA29152;
 XX
 XX
 DT 12-SEP-2000 (first entry)
 XX
 DE Partial gamma tocopherol methyltransferase cDNA, clone r10n.pK085.e11.
 XX
 KW Vitamin E; alpha-tocopherol; biosynthesis; enzyme; inhibitor; herbicide;
 KW gamma-tocopherol methyltransferase; ss.
 XX
 OS Oryza sativa.
 XX
 FH Key Location/Qualifiers
 FT 1.246 /tag- a
 FT CDS /product= gamma_tocopherol_methyltransferase
 FT /partial
 PN WO200032757-A2.
 XX
 PD 08-JUN-2000.
 XX
 PF 02-DEC-1999: 99WO-US28588.
 XX
 PR 03-DEC-1998: 98US-0110781.
 XX
 PA (DUPO) DU POINT DE MEMOURS & CO E I.
 XX
 PI Cahoon RE, Coughlan SJ, Miao G, Rafalski JA;
 XX
 DR MPI: 2000-412309/35.
 DR P-PSDB: AAY96460.
 XX
 PT Polynucleotide encoding a vitamin E biosynthetic enzyme polypeptide

PT such as gamma tocopherol methyltransferase, useful for synthesizing
PT vitamin E or as a target for design and discovery of herbicides

PS Claim 3: Page 50; 82pp; English.

CC AAA29151-69 encode vitamin E (alpha-tocopherol) biosynthetic enzymes or
CC their fragments. The cDNA clones were identified by BLAST searches based
CC on similarity to Synchocystis sp. and Arabidopsis thaliana vitamin E
CC biosynthetic enzymes. The enzymes are useful for synthesizing
CC plastoquinones or tocopherols (especially vitamin E). Vitamin E is
CC required for plant growth, therefore the enzymes may be used for the
CC discovery of new herbicides. The enzymes can be used in methods to
CC evaluate potential inhibitors, which may have use as herbicides.
CC Additionally, transgenic expression of, e.g. gamma-tocopherol
CC methyltransferase affords the ability to manipulate tocopherol levels as
CC desired for a particular application.

XX Sequence 521 BP; 156 A; 114 C; 119 G; 116 T; 16 other;

SO Query Match 10.6%; Score 125.8; DB 21; Length 521;

Best Local Similarity 70.1%; Pred. No. 2.9e-24; Indels 0; Gaps 0;

Matches 169; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

OY 633 CTGCAAAAGCTAGTTTGTGGAGAGTTAGCTCGGTAGCAGCACCAGTGCATTATTA 692

DB 5 CAGACAAAGCGAGTTTGTAAAGCAGCTGCAGCCGTCCGAGCTCTGGGGCGAATAA 64

OY 693 TAATAGTACATGCTGCCACAGGATCTTGCCCTGACGAAACAATCTTACATCATTGGG 752

DB 65 TCATTGTGACCTGCTGCAATAGGAACTCCAGCCATCCGAAAGAGTCCCTGAACCTGATG 124

OY 753 AGCAAGATCTCTTAAAGAAGATTGGATGATATTAACCTCCGCTGGTGCCTCACTT 812

DB 125 AGCTGAATCTCTGAAAAGATATGCGATGATATTAATCTCCGAGACTGGTCTCTCTT 184

OY 813 CTGATTATGTTAAGTTGCTCCCAATCCCTGTCACCTTCAGGACATCAAGTCAGAAATTGGT 872

DB 185 CTGATTATGTCAAAATTGCCGAGTCACATGCTCTTGAAGATATTAAGACACGATTTGGT 244

OY 873 C 873

DB 245 C 245

RESULT 15

ID AAA29162 standard; cDNA; 521 BP.

XX AAA29162;

DT 12-SEP-2000 (first entry)

DE Rice gamma-tocopherol methyltransferase cDNA clone rls72.pk0010.g3-3'

KW Vitamin E: alpha-tocopherol; biosynthesis; enzyme; inhibitor; herbicide;

KW gamma-tocopherol methyltransferase; ss.

OS Oryza sativa.

PN MO200032757-A2.

PD 08-JUN-2000.

PE 02-DEC-1999; 99WO-US28588.

PR 03-DEC-1998; 98US-0110781.

PA (DUPO) DU PONT DE NEMOURS & CO E I.

PI Cahoon RE, Coughlan SJ, Miao G, Rafalski JA;

DR WPI; 2000-412309/35.

PT Polynucleotide encoding a vitamin E biosynthetic enzyme polypeptide
PT such as gamma tocopherol methyltransferase, useful for synthesizing
PT vitamin E or as a target for design and discovery of herbicides

PS Claim 3: Page 63; 82pp; English.

CC AAA29151-69 encode vitamin E (alpha-tocopherol) biosynthetic enzymes or
CC their fragments. The cDNA clones were identified by BLAST searches based
CC on similarity to Synchocystis sp. and Arabidopsis thaliana vitamin E
CC biosynthetic enzymes. The enzymes are useful for synthesizing
CC plastoquinones or tocopherols (especially vitamin E). Vitamin E is
CC required for plant growth, therefore the enzymes may be used for the
CC discovery of new herbicides. The enzymes can be used in methods to
CC evaluate potential inhibitors, which may have use as herbicides.
CC Additionally, transgenic expression of, e.g. gamma-tocopherol
CC methyltransferase affords the ability to manipulate tocopherol levels as
CC desired for a particular application.

XX Note: The specification indicates that this sequence should encode the
XX protein given in AAY96470.

SO Query Match 10.6%; Score 125.8; DB 21; Length 521;

Best Local Similarity 70.1%; Pred. No. 2.9e-24; Indels 0; Gaps 0;

Matches 169; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

OY 633 CTGCAAAAGCTAGTTTGTGGAGAGTTAGCTCGGTAGCAGCACCAGTGCATTATTA 692

DB 5 CAGACAAAGCGAGTTTGTAAAGCAGCTGCAGCCGTCCGAGCTCTGGGGCGAATAA 64

OY 693 TAATAGTACATGCTGCCACAGGATCTTGCCCTGACGAAACAATCTTACATCATTGGG 752

DB 65 TCATTGTGACCTGCTGCAATAGGAACTCCAGCCATCCGAAAGAGTCCCTGAACCTGATG 124

OY 753 AGCAAGATCTCTTAAAGAAGATTGGATGATATTAACCTCCGCTGGTGCCTCACTT 812

DB 125 AGCTGAATCTCTGAAAAGATATGCGATGATATTAATCTCCGAGACTGGTCTCTCTT 184

OY 813 CTGATTATGTTAAGTTGCTCCCAATCCCTGTCACCTTCAGGACATCAAGTCAGAAATTGGT 872

DB 185 CTGATTATGTCAAAATTGCCGAGTCACATGCTCTTGAAGATATTAAGACACGATTTGGT 244

OY 873 C 873

DB 245 C 245

Search completed: March 18, 2003, 07:21:16
Job time : 1536 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 18, 2003, 06:53:09 ; Search time 294 Seconds
(without alignments)
1240.268 Million cell updates/sec

Title: US-09-857-613A-27
Perfect score: 1189
Sequence: 1 ggcacatgcccacccgtgtgta.....aaaaaaaaaaaaaaaaaaaaa 1189

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCUTS.COMB.seq: *
6: /cgn2_6/ptodata/1/ina/Backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76.6	6.4	957	US-09-382-906A-1	Sequence 1, Appl 1
2	51	4.3	7218	US-08-232-463-14	Sequence 14, Appl 1
3	42.4	3.6	1400	US-09-041-718-1	Sequence 1, Appl 1
4	40.6	3.4	436	US-09-397-787-257	Sequence 257, Appl 1
5	40.2	3.4	1447	US-09-443-041A-27	Sequence 27, Appl 1
6	39.4	3.3	3157	5198347-3	Patent No. 5198347
7	38.8	3.3	1508	US-09-039-046-1	Sequence 1, Appl 1
8	38.8	3.3	2706	US-09-463-238-2	Sequence 2, Appl 1
9	38.8	3.3	2806	US-09-463-238-1	Sequence 11, Appl 1
10	38.2	3.2	1736	US-09-182-816-22	Sequence 22, Appl 1
11	38.2	3.2	1736	US-09-471-528-22	Sequence 22, Appl 1
12	38.2	3.2	1736	US-09-471-528-22	Sequence 22, Appl 1
13	38.2	3.2	1736	US-09-471-528-22	Sequence 22, Appl 1
14	38.2	3.2	1736	US-09-634-530-24	Sequence 22, Appl 1
15	38.2	3.2	1736	US-09-634-530-24	Sequence 24, Appl 1
16	37.8	3.2	1558	US-08-455-550-7	Sequence 7, Appl 1
17	37.8	3.2	3080	US-09-099-041A-25	Sequence 25, Appl 1
18	37.8	3.2	3080	US-09-245-281-25	Sequence 25, Appl 1
19	37.8	3.2	3080	US-09-207-359B-25	Sequence 25, Appl 1
20	37.8	3.2	6200	US-09-439-923-1	Sequence 13, Appl 1
21	37.8	3.2	19124	US-08-487-826B-13	Sequence 13, Appl 1
22	37.6	3.2	3138	US-07-867-106-4	Sequence 4, Appl 1
23	37.4	3.1	1332	US-09-412-600B-1	Sequence 1, Appl 1
24	37.4	3.1	7941	US-09-816-703A-1	Sequence 1, Appl 1
25	37	3.1	289	US-09-007-005-17	Sequence 17, Appl 1
26	37	3.1	289	US-09-244-796-17	Sequence 17, Appl 1
27	37	3.1	1986	US-08-687-590-57	Sequence 57, Appl 1

c	28	37	3.1	5852	1	US-07-867-106-2	Sequence 2, Appl 1
	29	36.8	3.1	1123	4	US-09-152-060-15	Sequence 15, Appl 1
	30	36.8	3.1	1210	4	US-09-443-041A-29	Sequence 29, Appl 1
	31	36.8	3.1	2280	1	US-08-813-150-1	Sequence 1, Appl 1
	32	36.8	3.1	3792	2	US-08-992-334-1	Sequence 1, Appl 1
	33	36.8	3.1	3792	3	US-08-302-752-1	Sequence 1, Appl 1
	34	36.8	3.1	5234	2	US-08-992-334-2	Sequence 2, Appl 1
	35	36.8	3.1	5234	3	US-08-302-752-2	Sequence 2, Appl 1
	36	36.8	3.1	6722	2	US-08-992-334-3	Sequence 3, Appl 1
	37	36.8	3.1	6722	3	US-08-302-752-3	Sequence 3, Appl 1
	38	36.4	3.1	144	1	US-08-702-344-26	Sequence 26, Appl 1
	39	36.4	3.1	1104	4	US-09-009-816-1	Sequence 9, Appl 1
	40	36.4	3.1	1172	1	US-07-945-288-9	Sequence 9, Appl 1
	41	36.4	3.1	1172	1	US-08-462-831-9	Sequence 9, Appl 1
	42	36.4	3.1	1172	1	US-08-461-809-9	Sequence 9, Appl 1
	43	36.4	3.1	1172	1	US-08-461-441-9	Sequence 9, Appl 1
	44	36.4	3.0	1172	5	PCF-US93-08518-9	Sequence 9, Appl 1
	45	36.2	3.0	1622	2	US-09-014-969-16	Sequence 16, Appl 1

ALIGNMENTS

RESULT 1
US-09-382-906A-1
: Sequence 1, Application US/09382906A
: Patent No. 6448475
: GENERAL INFORMATION:
: APPLICANT: Dellapenna, Dean
: APPLICANT: Shintani, David
: TITLE OF INVENTION: Manipulation of Tocopherol Levels in Transgenic Plants
: FILE REFERENCE: 920905, 90032
: CURRENT APPLICATION NUMBER: US/09/382, 906A
: PRIOR FILING DATE: 1999-08-25
: PRIOR APPLICATION NUMBER: 60/097, 863
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1
: LENGTH: 957
: TYPE: DNA
: ORGANISM: Synechocystis PCC6803
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(954)
US-09-382-906A-1

Query Match 6.4%; Score 76.6; DB 4; Length 957;
Best Local Similarity 54.7%; Pred. No. 3e-12;
Matches 179; Conservative 0; Mismatches 139; Indels 9; Gaps 1;

OY	399	TACTGATGTTGGGTGGGATAGTGGCAGCTCTACATGCTGGCCCAAGAAATTTGAG	458
DB	290	TATTGATGTTGGGTGGGATAGTGGCAGCTCTACATGCTGGCCCAAGAAATTTGAG	349
OY	459	CAACACGTAGTACATCTGAGTCTTCAAGCTCAAGACCAAGCAATGCTCTG	518
DB	350	TTACGTAGTACATCTGAGTCTTCAAGCTCAAGACCAAGCAATGCTCTG	409
OY	519	CTGCTCAAGTATGCTGATAGTCTTCAAGTCTGAGTCTGAGACCTCTACAGAACAT	578
DB	410	CTCCGAGTGGTACGCGCCAG-----TTTGGGTTGAGAGATGCTATGGCTTGTCTT	460
OY	579	TCTGTAGCGGATCTGATCTGCTGCTGATGAGAGTGGAGACATTTGCTGACA	638
DB	461	TTCTGTAGCGGATCTGATCTGCTGCTGATGAGAGTGGAGACATTTGCTGACA	520
OY	639	AAGCTAAGTTTGTGAGAGTTAGCTGCGTAGGACGACCAAGTTCATTAATTAATAG	698
DB	521	AAGCTAAGTTTGTGAGAGTTAGCTGCGTAGGACGACGAGGAGGATTTCTGTGTGG	580
OY	699	TAACTGTGCTCCACAGGATTTGCTG	725


```
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(436)
OTHER INFORMATION: n = A,T,C or G
US-09-397-787-257
```

```
Query Match
Best Local Similarity 3.4%; Score 40.6; DB 4; Length 436;
Matches 100; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
```

```
QY 30 CAATCTATGATCCATCCACACGTTCCCAATCCCTTCGACCTTCCGACGAA 89
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 6 CACCCGCGCGGCTCCAGGGCCCTCCGCTCGGGCCCGGGACCCCGGCTCCGACG 65
QY 90 TCCGGGTCCGACCCAGGTCTGGGCTCTTTCGGGCATGGCAGCGACGTCGAGAG 149
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 66 CCGGCCCGCGCCCGCCGACCATGTCNGATAAAGCGTGAGAGCGACGCGAGTTGAGCG 125
QY 150 GGGAGATGATTTGAGAGCAGAGCCGAGATGACAGAGATGACAGAGGAA 209
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 126 CCAAGACCTGAAAGAGAGAGAGGTTGGAGAGAGCAGCCGGAAGAGGAA 185
QY 210 TCGCAGAGTTTACGACGAG 229
    ||| ||| ||| |||
DB 186 AGAAGAGATGCTGAGAGAG 205
```

RESULT 5

```
US-09-443-041A-27
Sequence 27, Application US/09443041A
Patent No. 6455717
GENERAL INFORMATION:
APPLICANT: Famodu, Omolayo O.
APPLICANT: Orozco, Buddy
APPLICANT: Rafalski, Antoni
APPLICANT: Shen, Jennie
TITLE OF INVENTION: Sterol Metabolism Enzymes
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/09/443,041A
CURRENT FILING DATE: 1999-11-18
PRIOR APPLICATION NUMBER: 60/109,283
PRIOR FILING DATE: 1998-11-20
NUMBER OF SEQ ID NOS: 33
SOFTWARE: Microsoft Office 97
SEQ ID NO 27
LENGTH: 1447
TYPE: DNA
ORGANISM: Glycine max
US-09-443-041A-27
```

```
Query Match
Best Local Similarity 3.4%; Score 40.2; DB 4; Length 1447;
Matches 69; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
```

```
QY 1073 TACTTTATGATGACCAAGTTCCACAGCTGCTTATTTGATGATGAGAAACAA 1132
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1229 TATGTTTATGAAATTAACCAACACTCTCTAAGGGTGTGTTGTTGTTAAAGAA 1288
QY 1133 GAGAAATAATTAATGAAGGGTGTGATTTTAAATAAAAAAAAAAAAAAAAAA 1189
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1289 TGGGAATGCAAGAGAAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1345
```

RESULT 6

```
5198347-3
Patent No. 5198347
APPLICANT: MILLER, LOUIS H.; ADAMS, JOHN H.; KASLOW,
DAVID C.; FANG, XIANGDOUG
TITLE OF INVENTION: DNA ENCODING PLASMODIUM VIVAX AND
PLASMODIUM KNOWLESII DUFFY RECEPTOR
NUMBER OF SEQUENCES: 27
```

```
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/554,837
FILING DATE: 20-JUL-1990
SEQ ID NO: 3
LENGTH: 3157
5198347-3
```

```
Query Match
Best Local Similarity 3.3%; Score 39.4; DB 6; Length 3157;
Matches 55; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
```

```
QY 1109 TTTATTTGATGTTGAGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1168
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3076 TATTTTCTATATTTGACAAATAAAAAAAAAAAAAAAAAATAACATTCAATATTTGA 3135
QY 1169 AAAAAAAAAAAAAAAAAAAAAA 1189
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3136 AAAAAAAAAAAAAAAAAAAAAA 3156
```

RESULT 7

```
US-09-039-046-1
Sequence 1, Application US/09039046
Patent No. 6331660
GENERAL INFORMATION:
APPLICANT: CHOMET, PAUL S.
APPLICANT: FREY, MONIKA
APPLICANT: GIERL, ALFONS
TITLE OF INVENTION: MAIZE DIMBOA BIOSYNTHESIS GENES
FILE REFERENCE: DKM:131
CURRENT APPLICATION NUMBER: US/09/039,046
CURRENT FILING DATE: 1998-03-13
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 1
LENGTH: 1508
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: CDS
LOCATION: (49)..(1089)
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: Primer
US-09-039-046-1
```

```
Query Match
Best Local Similarity 3.3%; Score 38.8; DB 4; Length 1508;
Matches 67; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
```

```
QY 1076 TTTATGATGATGACCAAGTTCCACAGCTGCTTATTTGATGATGAGAAACAGAG 1135
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1370 TCTGTGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1429
QY 1136 AAAAAAAAAATAATGAAGGGTGTGATTTTAAATAAAAAAAAAAAAAAAAAA 1189
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1430 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1483
```

RESULT 8

```
US-09-463-238-2
Sequence 2, Application US/09463238
Patent No. 6469230
GENERAL INFORMATION:
APPLICANT: Edwards, Elizabeth A
APPLICANT: Smith, Alison M
APPLICANT: Bustos Guillen, Regla
APPLICANT: Martin, Catherine R
APPLICANT: Plant Bioscience Limited
TITLE OF INVENTION: Starch Debranching Enzymes
FILE REFERENCE: 97.118
CURRENT APPLICATION NUMBER: US/09/463,238
CURRENT FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: PCT/GB98/02280
```

;; PRIOR FILING DATE: 1998-07-30
;; PRIOR APPLICATION NUMBER: GB 9716185.5
;; PRIOR FILING DATE: 1997-07-31
;; NUMBER OF SEQ ID NOS: 30
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 2
;; LENGTH: 2706
;; TYPE: DNA
;; ORGANISM: Solanum tuberosum
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (2641)
;; OTHER INFORMATION: n = a or g or c or t
US-09-463-238-2

Query Match
Best Local Similarity 66.3%; Score 38.8; DB 4; Length 2706;
Matches 55; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

OY 1107 CGTTTATTCGATGAGTGGAGAACAGAGAAAAATTAATGAAGGGTTGTCGATTTT 1166
DB 2606 CTGTATTTTAAATTCGATTTATATAAATAATANTAGTGAATTTGCTGCGAAAAA 2665
OY 1167 AAAAAAAAAAAAAAAAAAAAAA 1189
DB 2666 AAAAAAAAAAAAAAAAAAAAAA 2688

RESULT 9

US-09-463-238-11
;; Sequence 11, Application US/09463238
;; Patent No. 6469230
;; GENERAL INFORMATION:
;; APPLICANT: Edwards, Elizabeth A
;; APPLICANT: Smith, Alison M
;; APPLICANT: Bustos Guillen, Regla
;; APPLICANT: Martin, Catherine R
;; APPLICANT: Plant Bioscience Limited
;; TITLE OF INVENTION: Starch Dextranizing Enzymes
;; FILE REFERENCE: 97.118
;; CURRENT APPLICATION NUMBER: US/09/463,238
;; PRIOR FILING DATE: 2000-01-21
;; PRIOR APPLICATION NUMBER: PCT/GB98/02280
;; PRIOR FILING DATE: 1998-07-30
;; PRIOR APPLICATION NUMBER: GB 9716185.5
;; PRIOR FILING DATE: 1997-07-31
;; NUMBER OF SEQ ID NOS: 30
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 11
;; LENGTH: 2806
;; TYPE: DNA
;; ORGANISM: Solanum tuberosum
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (822, 826, 2707 and 2797)
;; OTHER INFORMATION: n = a or g or c or t
US-09-463-238-11

Query Match
Best Local Similarity 66.3%; Score 38.8; DB 4; Length 2806;
Matches 55; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

OY 1107 CGTTTATTCGATGAGTGGAGAACAGAGAAAAATTAATGAAGGGTTGTCGATTTT 1166
DB 2672 CTGTATTTTAAATTCGATTTATATAAATAATANTAGTGAATTTGCTGCGAAAAA 2731
OY 1167 AAAAAAAAAAAAAAAAAAAAAA 1189
DB 2732 AAAAAAAAAAAAAAAAAAAAAA 2754

RESULT 10
US-09-182-816-22

;; Sequence 22, Application US/09182816
;; Patent No. 6143542
;; GENERAL INFORMATION:
;; APPLICANT: Wisniewski, Nancy
;; APPLICANT: Silver, Gary M.
;; APPLICANT: Lo, Katherine C.
;; APPLICANT: Brandt, Kevin S.
;; TITLE OF INVENTION: NOVEL FLEA EPOXIDE HYDROLASE NUCLEIC ACID MOLECULES,
;; TITLE OF INVENTION: PROTEINS AND USES THEREOF
;; FILE REFERENCE: FC-3-C1
;; CURRENT APPLICATION NUMBER: US/09/182,816
;; EARLIER APPLICATION NUMBER: 08/989,510
;; EARLIER FILING DATE: 1997-12-12
;; NUMBER OF SEQ ID NOS: 31
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 22
;; LENGTH: 1736
;; TYPE: DNA
;; ORGANISM: Ctenocephalides felis
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (159)..(1553)
US-09-182-816-22

Query Match
Best Local Similarity 50.3%; Score 38.2; DB 3; Length 1736;
Matches 94; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

OY 1003 ATACAGAGAAGATCTAATTAATGTTGCCATCATTCAGTGCAGAAACCTGAATAATGA 1062
DB 1538 AGAAGAAATGAGATCTCGAAGAACTGTAAATTAATTTGTGAATAATAATATGTTAA 1597
OY 1063 GAGCAGAGATTACTTTATPAGAAATGACCAAGTTTCCAGAGCGTTTATTCGATAGT 1122
DB 1598 AATTAATGTAAATTAATGTAATTAACGATTAATGATTTTATTCGAACCTGCAAAAT 1657
OY 1123 TGAGAAACAGAGAAAAATTAATGAAGGGTTGTCGATTTTAAAAA 1182
DB 1658 AAAAAAAAAAAAAAAAAAAAAA 1717
OY 1183 AAAAAA 1189
DB 1718 AAAAAA 1724

RESULT 11

US-09-182-816-24/C
;; Sequence 24, Application US/09182816
;; Patent No. 6143542
;; GENERAL INFORMATION:
;; APPLICANT: Wisniewski, Nancy
;; APPLICANT: Silver, Gary M.
;; APPLICANT: Lo, Katherine C.
;; APPLICANT: Brandt, Kevin S.
;; TITLE OF INVENTION: NOVEL FLEA EPOXIDE HYDROLASE NUCLEIC ACID MOLECULES,
;; TITLE OF INVENTION: PROTEINS AND USES THEREOF
;; FILE REFERENCE: FC-3-C1
;; CURRENT APPLICATION NUMBER: US/09/182,816
;; EARLIER APPLICATION NUMBER: 08/989,510
;; EARLIER FILING DATE: 1997-12-12
;; NUMBER OF SEQ ID NOS: 31
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 24
;; LENGTH: 1736
;; TYPE: DNA
;; ORGANISM: Ctenocephalides felis
US-09-182-816-24

Query Match
Best Local Similarity 50.3%; Score 38.2; DB 3; Length 1736;
Matches 94; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

RESULT 13
US-09-471-528-24/C
Sequence 24, Application US/09471528
Patent No. 6153397
GENERAL INFORMATION:
APPLICANT: Wisniewski, Nancy

Query Match	Score	38.2	DB	4	Length	1736			
Best Local Similarity	50.38	Pred	No	0.33					
Matches	94	Conservative	0	Mismatches	93	Indels	0	Gaps	0

```
OY 1003 ATACAGAGATCTAATTAGTTGGCATTCATGTCGAACCAACCTGATAATGGA 1062
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1538 AGAGAAATGAGATCTCTGAGAACTGTGAATTAATTTGTGATAATATATATGTTAA 1597
OY 1063 GAGCAGAGATTACTTTTATAGATGAACCAAGTTCCACAGTGTGTTATTCGATAGT 1122
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1598 AAATAAATGTATTAATGCTGAATTAACGATATGATTTTATTCAAACTGTCAATAT 1657
OY 1123 TGAGAAACAAGAGAAAAAATTAATGAAGGGGTTGTCGATTTTAAAAA 1182
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1658 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1717
OY 1183 AAAAAA 1189
      | | | | | | |
Db 1718 AAAAAA 1724
```

```
RESULT 15
US-09-634-530-24/C
; Sequence 24, Application US/09634530
; Patent No. 6290958
; GENERAL INFORMATION:
; APPLICANT: Wisniewski, Nancy
; APPLICANT: Silver, Gary M.
; APPLICANT: Lo, Katherine C.
; APPLICANT: Brandt, Kevin S.
; TITLE OF INVENTION: FLEA EPOXIDE HYDROLASE PROTEINS AND USES THEREOF
; FILE REFERENCE: FC-3-CI-1
; CURRENT APPLICATION NUMBER: US/09/634,530
; CURRENT FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 09/471,528
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 09/182,816
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 08/989,510
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentl Ver. 2.0
; SEQ ID NO 24
; LENGTH: 1736
; TYPE: DNA
; ORGANISM: Ctenocephalides felis
US-09-634-530-24
```

```
Query Match 3.28; Score 38.2; DB 4; Length 1736;
Best Local Similarity 50.34; Pred. No. 0.33;
Matches 94; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
```

```
OY 1003 ATACAGAGATCTAATTAGTTGGCATTCATGTCGAACCAACCTGATAATGGA 1062
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 199 AGAGAAATGAGATCTCTGAGAACTGTGAATTAATTTGTGATAATATATATGTTAA 140
OY 1063 GAGCAGAGATTACTTTTATAGATGAACCAAGTTCCACAGTGTGTTATTCGATAGT 1122
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 139 AAATAAATGTATTAATGCTGAATTAACGATATGATTTTATTCAAACTGTCAATAT 80
OY 1123 TGAGAAACAAGAGAAAAAATTAATGAAGGGGTTGTCGATTTTAAAAA 1182
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 79 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 20
OY 1183 AAAAAA 1189
      | | | | | | |
Db 19 AAAAAA 13
```

Search completed: March 18, 2003, 11:33:02
Job time : 317 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: March 18, 2003, 08:23:37 ; Search time 108 Seconds
(without alignments)
7727.015 Million cell updates/sec

Title: US-09-857-613a-27

Perfect score: 1189

Sequence: 1 ggacatggccaccctggtgta.....aaaaaaaaaaaaaaaaaaaaa 1189

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 501302 seqs, 350932545 residues

Total number of hits satisfying chosen parameters: 1002604

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

ALIGNMENTS

```
20 40.4 3.4 2394 10 US-09-800-729-33 Sequence 33, Appl
21 40.2 3.4 2000 9 US-09-938-842A-3966 Sequence 3966, Ap
22 40.2 3.4 3313 10 US-09-764-846-83 Sequence 83, Appl
23 40.2 3.4 225 9 US-09-925-299-744 Sequence 744, Ap
24 40.2 3.4 225 10 US-09-925-299-744 Sequence 744, Ap
25 40.2 3.4 225 10 US-09-960-352-5879 Sequence 5879, Ap
26 39.8 3.3 446 10 US-09-960-352-3400 Sequence 3400, Ap
27 39.6 3.3 361 9 US-10-125-258-108 Sequence 108, Ap
28 39.6 3.3 681 10 US-09-822-830A-26 Sequence 26, Appl
29 39.6 3.3 166 10 US-09-867-701-9636 Sequence 9636, Ap
30 39.4 3.3 302 9 US-09-925-299-710 Sequence 710, Ap
31 39.4 3.3 302 10 US-09-925-299-710 Sequence 710, Ap
32 39.4 3.3 1290 10 US-09-764-846-124 Sequence 124, Ap
33 39.2 3.3 227 10 US-09-878-574-11011 Sequence 11011, A
34 39.2 3.3 239 10 US-09-960-352-11438 Sequence 11438, A
35 39 3.3 3819 9 US-09-978-295A-482 Sequence 482, Ap
36 39 3.3 3819 9 US-09-978-295A-482 Sequence 482, Ap
37 39 3.3 3819 9 US-09-978-295A-482 Sequence 482, Ap
38 39 3.3 3819 9 US-09-978-295A-482 Sequence 482, Ap
39 39 3.3 3819 9 US-09-978-295A-482 Sequence 482, Ap
40 39 3.3 3819 9 US-09-978-295A-482 Sequence 482, Ap
41 39 3.3 3819 9 US-09-978-295A-482 Sequence 482, Ap
42 39 3.3 3819 9 US-09-978-295A-482 Sequence 482, Ap
43 39 3.3 3819 9 US-09-978-295A-482 Sequence 482, Ap
44 39 3.3 3819 9 US-09-978-295A-482 Sequence 482, Ap
45 39 3.3 3819 9 US-09-978-295A-482 Sequence 482, Ap
```

Database: 1: Published Applications - NA: *

```
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PC7_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PC7US_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	194.4	16.3	252	10 US-09-878-574-695	Sequence 695, App
2	124.8	10.5	311	10 US-09-294-0938-3472	Sequence 3472, Ap
3	44.6	3.8	286	10 US-09-960-352-13342	Sequence 13342, A
4	43.6	3.7	2045	10 US-09-925-300-749	Sequence 749, App
5	42.6	3.6	932	12 US-10-078-929-65	Sequence 65, Appl
6	42.2	3.5	368	10 US-09-834-975-47	Sequence 47, Appl
7	42.2	3.5	2000	9 US-09-938-842A-3589	Sequence 3589, Ap
8	41.8	3.5	401	10 US-09-960-352-14060	Sequence 14060, A
9	41.6	3.5	1046	10 US-09-925-297-307	Sequence 307, App
10	41.4	3.5	393	10 US-09-960-352-4582	Sequence 4582, Ap
11	41.2	3.5	620	10 US-09-925-297-303	Sequence 303, App
12	41.2	3.5	4840	10 US-09-880-107-3423	Sequence 3423, Ap
13	41	3.4	347	10 US-09-960-352-3371	Sequence 3371, Ap
14	41	3.4	419	10 US-09-960-352-11234	Sequence 11234, A
15	41	3.4	2608	10 US-09-739-254-16	Sequence 16, Appl
16	41	3.4	2608	10 US-09-904-615-16	Sequence 16, Appl
17	40.8	3.4	1457	9 US-09-989-920-93	Sequence 93, Appl
18	40.6	3.4	436	10 US-09-876-889-257	Sequence 257, App
19	40.4	3.4	2059	9 US-09-957-708-4	Sequence 4, Appl

```
RESULT 1
US-09-878-574-695
Sequence 695, Application US/09878574
Patent No. US20020110548A1
GENERAL INFORMATION:
APPLICANT: Byrum, Joseph R.
APPLICANT: La Rosa, Thomas J.
APPLICANT: Thompson, Michael D.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated w/
TITLE OF INVENTION: Plants
FILE REFERENCE: 38-21(15401)B
CURRENT APPLICATION NUMBER: US/09/878,574
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 09/333,535
PRIOR FILING DATE: 1999-06-14
NUMBER OF SEQ ID NOS: 15775
SEQ ID NO 695
LENGTH: 252
TYPE: DNA
ORGANISM: Glycine max
OTHER INFORMATION: Clone ID: LIB3028-047-Q1-B1-D12
US-09-878-574-695
```

```
Query Match 16.3%, Score 194.4; DB 10; Length 252;
Best Local Similarity 88.5%; Pred. No. 6e-40;
Matches 223; Conservative 0; Mismatches 26; Indels 3; Gaps 1;
122 AGAAGCTGACAGAGGATGCGAGGTTTTCGAGAGGTCGTCGTTATGGAGACA 251
1 AAAATGTCGAGAGGAAATTCGAGATTCGAGAGGTCGTCGTTATGGAGACA 60
252 TTGGGGGACACATGACATGACATGCTTTATGATCGATTCATCTTTCGCTTCG 311
TTTGGGGGACATGACATGACATGACATGCTTTATGATCGATTCATCTTTCGCTTCG 120
61 TTTGGGGGACATGACATGACATGACATGCTTTATGATCGATTCATCTTTCGCTTCG 120
312 ATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 368
121 ATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
369 CTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 428
181 CTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
```

Oy	429	GCTCTAGATACC	440
Db	241	GCTCAGATACC	252

RESULT 2

Sequence 3472, Application US/09294093B
Patent No. US20010051335A1
GENERAL INFORMATION:
APPLICANT: Ialugudi, Raghunath, V.
APPLICANT: Ito, laura, Y.
APPLICANT: Sherman, Bradley, K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
FIR REPRESENTOR: R 0000 US

```

? SEQ ID NO 3472
? LENGTH: 311
? TYPE: DNA
? ORGANISM: Zea mays
? FEATURE:
? NAME/KEY: misc_feature
? OTHER INFORMATION: Incyte ID No. US20010051335A1 700380165H1
? LOCATION: 4, 43, 55, 96, 99, 127, 156
? OTHER INFORMATION: a, t, c, g, or other
? S-09-294 -093B-3472

```

Query Match	10.5%	Score 124.8	DB 10	length 311
Best Local Similarity	63.3%	Pred. No. 2.9e-22		
Matches 186	Conservative	0	Mismatches 108	Indels 0
				Gaps 0

QY	732	AACAAATCCCTTAACATCCATGTGGAGCAACATCTCTTAAAGAAGATTGGAGTCATATATAC	79
Db	5	AACCTGCCTAAAGCCCGCATGACCTGACCTCTCTGAAANANATATGGCGAANCGTACTAC	64
QY	792	TCCCTGCCTGTGCTCAACTTCTGATTATGTTAAATTGCTCAATGCCCTGTCACTCAGG	85
Db	65	TCCCGGACTGTGCTCACTTCAACACTATGTAAANATTTGCCAAATCACTGCTCTCGAG	124
QY	852	ACATCAAGTCAGAAAGATTGGCTCTCGCTTTGTCATTTTGGCCAGCAGTGTACGCT	91
Db	125	ATATCAACGACAGCTGACTGTGCGAGAACGTCGCCCCGTTTGGCCCGCGCTGTAAAT	188
QY	912	CAGCCTTCACATGGAAGGGTCTATCTTCACCTCTTGACACAGTGCACAAAACGATTAAG	97
Db	185	CAGCGCTAAACATGGAAGGGCTTCACTCTCTGCTGACGACCCGATGGAAGCATTCAGAG	244
QY	972	GAGCTTGGCTATGCCATTGATGATTAAGGGATTCAAAGAAAGATCTTAATTAAT	1025
Db	245	GCGGGAATGGATGCCCTATATGATTCACAGGGCTTCAAGAAAGGGCTCTATCAATT	298

RESULT 3

```

: Sequence 13342, Application US/09960352
: Patent No. US20020137139A1
:
: GENERAL INFORMATION:
:
: APPLICANT: Warren, Wesley C.
: APPLICANT: Tao, Nengbing
: APPLICANT: Byatt, John C.
: APPLICANT: Mathialagan, Nagappan
:
: TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
: FILE REFERENCE: 16511.006/37-21(102981C
: CURRENT APPLICATION NUMBER: US/09/960,352
: CURRENT FILING DATE: 2001-09-24

```

```

? NUMBER OF SEQ ID NOS: 15112
? SEQ ID NO 13342
? LENGTH: 286
? TYPE: DNA
? ORGANISM: Bos taurus
? OTHER INFORMATION: Clone ID: 57-LIB3058-009-Q1-K1-G2
? OS-09-960-352-13342

```

Query Match	3.88;	Score 44.6;	DB 10;	Length 286;
Best Local Similarity	57.68;	Pred. No. 0.061;		
Matches 80; Conservative	0;	Mismatches 59;	Indels 0;	Gaps 0;

QY	1051	TGAATTAATGAGAGGCGAGATTACTTTATATGAAAGAAACCAAGTTCACAGCGTCGT	1110
Db	272	TTATATAATTAAGAAAGCAAAAAGTTTATATAGCCAAATTAAGCAATAAAAAATATAA	213
QY	1111	TATTCGATGTTGAGACACAAAGAAAATTAATGAAAGGGTTCGATTTTAAA	1170
Db	212	TATTAGCCCATTAATAAATATTTTAAAAAAGAAAGTTTCTTACTTAAAAAAA	153
QY	1171	AAAAAAAAAAAAAAAAAAAAA	1189
Db	152	AAAAAAAAAAAAAAAAAAAAA	134

RESULT 4

```

? Sequence 749, Application US/09925300
? Patent No. US20020151681A1
? GENERAL INFORMATION:
? APPLICANT: Craig Rosen,
? APPLICANT: Steve Ruben
? TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
? FILE REFERENCE: PA101
? CURRENT APPLICATION NUMBER: US/09/925,300
? PRIOR FILING DATE: 2001-08-10
? PRIOR APPLICATION NUMBER: PCT/US00/05988
? PRIOR FILING DATE: 2000-03-08
? PRIOR APPLICATION NUMBER: 60/124,270
? PRIOR FILING DATE: 1999-03-12
? NUMBER OF SEQ ID NOS: 1890
? SOFTWARE: patentIn Ver. 2.0
? SEQ ID NO 749
? LENGTH: 2045
? TYPE: DNA
? ORGANISM: Homo sapiens
? US-09-925-300-749

```

Query Match	3.78;	Score 43.6;	DB 10;	Length 2045;
Best Local Similarity	60.98;	Pred. No. 0.27;		
Matches 67; Conservative	2;	Mismatches 41;	Indels 0;	Gaps 0;

QY	1080	1929	1140	1989
ATGAAATGAAACCAAGTTCCACAGGTCGTTATTTCCGATCTGAGAACCAAGGAAA	AAATPAAGSACCTCTTTTAATACCAAGCATTTGTACTGTTCTTTTATATACGAAA	AATPAATGAAAGGCGTTGTCATTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAA	TATPAAAAAGGAGGKGTGTGTTAAAAAARAAAAAAAAAAAAAAAAAAAAAA	
1139	1988	1189	2038	

RESULT. 5

US-10-07-8-925-63
Sequence 65, Application US/10078923
Patent No. US20020152497A1
GENERAL INFORMATION:
APPLICANT: Rafalski, Antoni
APPLICANT: Miao, Guo-Hua
APPLICANT: Falco, Saverio Carl
APPLICANT: Sakai, Hajime
APPLICANT: Farnoud, Omolayo O.
APPLICANT: Odell, Joan T.
APPLICANT: Meyers, Blake


```

? RESULT 6
? US-09-834-975-47/c
? Sequence 47, Application US/09834975
? Patent No. US20020110815A1
?
? GENERAL INFORMATION:
?
? APPLICANT: Lillie, James
? APPLICANT: Brown, Jeffrey
? APPLICANT: Bolt, Andrew
? APPLICANT: Van Huffel, Christophe
?
? TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
? TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
?
? TITLE OF INVENTION: OF HUMAN CANCERS
?
? FILE REFERENCE: MRL-016B
?
? CURRENT APPLICATION NUMBER: US/09/834, 975
? CURRENT FILING DATE: 2001-04-13
? PRIOR APPLICATION NUMBER: 60/197,538
? PRIOR FILING DATE: 2000-04-14
? NUMBER OF SEQ ID NOS: 1046
?
? SOFTWARE: FastSeq for Windows Version 4.0
?
? SEQ ID NO 47
?
? LENGTH: 368
?
? TYPE: DNA
?
? ORGANISM: Homo sapiens
?

```

```

      RESULT 7
      US-09-938-842A-3589
      ? Sequence 3589, Application US/09938842A
      ? Patent No. US20020160378A1
      ? GENERAL INFORMATION:
      ? APPLICANT: Harper, Jeff
      ? APPLICANT: Kreps, Joel
      ? APPLICANT: Wang, Xun
      ? APPLICANT: Zhu, Tong
      ? TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
      ? FILE REFERENCE: SAME, AND METHODS OF USE
      ? FILE REFERENCE: SCRIPT300-3
      ? CURRENT APPLICATION NUMBER: US/09/938,842A
      ? CURRENT FILING DATE: 2001-08-24
      ? PRIOR APPLICATION NUMBER: US 60/227,866
      ? PRIOR FILING DATE: 2000-08-24
      ? PRIOR APPLICATION NUMBER: US 60/264,647
      ? PRIOR FILING DATE: 2001-01-16
      ? PRIOR APPLICATION NUMBER: US 60/300,111
      ? PRIOR FILING DATE: 2001-06-22
      ? NUMBER OF SEQ ID NOS: 5379
      ? SEQ ID NO 3589
      ? LENGTH: 2000
      ? TYPE: DNA
      ? ORGANISM: Arabidopsis thaliana
      US-09-938-842A-3589

```

[illegible]

```
RESULT 8
US-09-960-352-14060/c
; Sequence 14060, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960.352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 14060
; LENGTH: 401
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 60-LIB3058-020-Q1-K1-G12
US-09-960-352-14060

Query Match          3.5%; Score 41.8; DB 10; Length 401;
Best Local Similarity 48.5%; Pred. No. 0.36;
Matches 115; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

OY 953 GGACAAAAAAGATAAAGAGCTTGGCTATGCCATTGATAGAGAGATACAGAAA 1012
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 264 GAATAAAGGAATTTAAAGGGGTTTAAAGTCCCAAGGTGCAAGGGTTGCAAAA 205
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1013 GATTAATTAAGTTTGCATCATACATGTCGAAAACCTGATAAATGAGAGCAGCAT 1072
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 204 AAAAAAAGATTTAAAGAAAAAAGAGAAAGTTGAGATAGAGTGCAGGGGTTT 145
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1073 TACTTTATGATGACCAAGCTTTCACACAGTCGTTTATTTGATAGTTGAGAAACA 1132
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 144 AACCAAGTTTGACAAAAAATTTAAGAAATTAAGCGAATTTTACCTGAAAAA 85
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1133 GAGAAAAAATTAATGAAGGGGTTTCGATTTTAAAAA 1189
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 84 AAAAAAAGATTTAAAGAAAAAAGAGAAAGTTGAGATAGAGTGCAGGGGTTT 28
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 9
US-09-925-297-307
; Sequence 307, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925.297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: Patentl Ver. 2.0
; SEQ ID NO 307
; LENGTH: 1046
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (14)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
```

```
; LOCATION: (946)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (946)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-297-307

Query Match          3.5%; Score 41.6; DB 10; Length 1046;
Best Local Similarity 51.7%; Pred. No. 0.63;
Matches 89; Conservative 1; Mismatches 82; Indels 0; Gaps 0;

OY 1018 AATTAAGTTTCCCATCATATACATGCGAAAACCTGATAAATGAGAGCAGATTACTT 1077
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 845 AAGTGATGACCTTTTTCGATGAAATCTAAATCTTTGTAACCTTTTATATCTG 904
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1078 TTAATGAATGACCAAGTTTCCACAGGTCGTTTATTTGATGATGAGAAACAAGAA 1137
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 905 CTTTGTTCACCAAGAAACCTAAATCCTTTTAMNANANAAAAA 964
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1138 AAATTAATGAAGGGTTCGATTTAAAAA 1189
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 965 AAAAAAAGATTTAAAGAAAAAAGAGAAAGTTGAGATAGAGTGCAGGGGTTT 1016
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 10
US-09-960-352-4582/c
; Sequence 4582, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960.352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 4582
; LENGTH: 393
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 20-LIB3057-013-Q1-K1-E7
US-09-960-352-4582

Query Match          3.5%; Score 41.4; DB 10; Length 393;
Best Local Similarity 48.5%; Pred. No. 0.46;
Matches 114; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

OY 955 ACAAAAAAGATTAAGAGAGCTTGGCTATGCCATGATGATAGAGGATACAGAAAGA 1014
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 352 AAAAAAATTTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 293
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1015 TCTAATTAAGTTTGCATCATATACATGTCGAAAACCTGATAAATGAGAGCAGGATTA 1074
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 292 TAATTAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 233
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1075 CTTTATGAATGACCAAGTTTCCACAGGTCGTTTATTTGATGATGAGAAACAAGAA 1134
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 232 TATAATTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 173
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1135 GAAAAAATTAATGAAGGGGTTTCGATTTTAAAAA 1189
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 172 AAAAAAAGATTTAAAGAAAAAAGAGAAAGTTGAGATAGAGTGCAGGGGTTT 118
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 11
US-09-925-297-303
; Sequence 303, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
```

Query Match	3.48;	Score 41;	DB 10;	Length 419;
Best Local Similarity	48.5%;	Pred. No. 0.59;		
Matches 113; Conservative	0;	Mismatches 120;	Indels 0;	Gaps

Q7 957 AAAAAACGATATAAAGGAGCTTTGGCGATTGCGCATGATGATAGAGGCGATCAAGAAAGATC 1016

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 17:36:17 ; Search time 38 Seconds
(without alignments)
1227.308 Million cell updates/sec

Title: US-09-857-613A-28

Perfect score: 1830
Sequence: 1 MATVVRIPITSCIHHTFPRS.....IEGYKKDLIKFATITCKRPE 350

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_101002:*

1:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1830	100.0	350	21	AA196472	Soybean gamma-toco
2	1668.5	91.2	349	21	AA196462	Soybean gamma-toco
3	1173	64.1	366	21	AA196473	Soybean gamma-toco
4	1162	63.5	352	21	AA196469	Corn gamma-tocophe
5	1154.5	63.1	348	21	AA196469	Arabidopsis thalia
6	1154.5	63.1	348	23	AB891452	Herbicidally activ
7	1151.5	62.9	308	21	AA196469	Arabidopsis thalia
8	1151.5	62.9	348	20	AA196469	Arabidopsis thalia
9	1151.5	62.9	348	21	AA196469	A. thaliana enviro
10	1151.5	62.9	348	23	AA196469	A. thaliana gamma-t

11	1147.5	62.7	348	21	AA196469	Arabidopsis thali
12	1144.5	62.5	308	21	AA196469	Arabidopsis thali
13	1035.5	56.6	266	21	AA196469	Arabidopsis thali
14	1028.5	56.2	266	21	AA196469	Arabidopsis thali
15	1022.5	55.9	266	22	AA196469	Physcomitrella th
16	944	51.6	293	21	AA196469	Partial wheat gam
17	778	42.5	191	21	AA196469	Corn gamma-tocoph
18	673	36.8	172	21	AA196469	Rice gamma-tocoph
19	647.5	35.4	317	20	AA196469	Synchocystis gam
20	392	21.4	301	22	AA196469	Synchocystis PCC
21	392	21.4	318	21	AA196469	Amino acid sequen
22	392	21.4	318	22	AA196469	Synchocystis PCC
23	392	21.4	318	22	AA196469	Synchocystis PCC
24	392	21.4	318	23	AA196469	Synchocystis 2-m
25	315	17.2	82	21	AA196469	Partial rice gamma
26	280.5	15.3	283	21	AA196469	S. lavenulae Mit
27	265	14.5	81	22	AA196469	Physcomitrella pa
28	257.5	14.1	128	21	AA196469	Rice gamma-tocoph
29	254	13.9	280	18	AA196469	S-adenosyl methio
30	232.5	12.7	285	21	AA196469	A. halophilla sarc
31	227	12.4	344	21	AA196469	Maize C-24 sterol
32	226	12.3	279	21	AA196469	E. halochloris sa
33	226	12.3	359	21	AA196469	Arabidopsis thali
34	226	12.3	359	23	AA196469	Herbicidally acti
35	224	12.2	144	21	AA196469	Partial rice gamma
36	222	12.1	115	21	AA196469	SAM dependent met
37	220.5	12.0	275	20	AA196469	Spnf protein invo
38	220.5	12.0	275	22	AA196469	S. spinosa protei
39	217	11.9	344	20	AA196469	Corn SMR enzyme
40	212.5	11.6	361	20	AA196469	Arabidopsis SMR e
41	212.5	11.6	361	23	AA196469	Herbicidally acti
42	207.5	11.3	346	23	AA196469	Nicotiana tabacum
43	203.5	11.1	293	21	AA196469	Arabidopsis thali
44	203.5	11.1	293	21	AA196469	Arabidopsis thali
45	203.5	11.1	293	21	AA196469	Arabidopsis thali

ALIGNMENTS

RESULT 1	AA196472	standard; Protein: 350 AA.
ID	AA196472	
AC	AA196472	
XX	12-SEP-2000	(first entry)
DE	Soybean gamma-tocopherol methyltransferase (clone sahlc.pK004.g2).	
XX	Vitamin E: alpha-tocopherol; biosynthesis; enzyme; inhibitor; herbicide;	
KW	gamma-tocopherol methyltransferase.	
XX	Glycine max.	
OS	MO200032757-A2.	
PN	08-JUN-2000.	
XX	02-DEC-1999;	99WO-US28588.
PD	03-DEC-1998;	98US-0110781.
XX	(DUPLO) DU PONT DE NEMOURS & CO E I.	
XX	Caboon RE, Coughlan SJ, Miao G, Rafalski JA;	
PI	WPI: 2000-412309/35.	
XX	N-PSDB: AKA29164.	
DR	Polynucleotide encoding a vitamin E biosynthetic enzyme polypeptide	
PT	such as gamma tocopherol methyltransferase, useful for synthesizing	
XX	vitamin E or as a target for design and discovery of herbicides	

XX Claim 11; Page 65-66; 82pp; English.

PS AAY96459-77 encode vitamin E (alpha-tocopherol) biosynthetic enzymes or

CC their fragments. The cDNA clones were identified by BLAST searches based

CC on similarity to Synecocystis sp. and Arabidopsis thaliana vitamin E

CC biosynthetic enzymes. The enzymes are useful for synthesizing

CC plastoquinones or tocopherols (especially vitamin E). Vitamin E is

CC required for plant growth, therefore the enzymes may be used for the

CC discovery of new herbicides. The enzymes can be used in methods to

CC evaluate potential inhibitors, which may have use as herbicides.

CC Additionally, transgenic expression of, e.g. gamma-tocopherol

CC methyltransferase affords the ability to manipulate tocopherol levels as

CC desired for a particular application.

CC

XX Sequence 350 AA;

SO

Query Match 100.0%; Score 1830; DB 21; Length 350;

Best Local Similarity 100.0%; Pred. No. 2e-172;

Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATVRIPTISCIHHTFRSOSPRTFARIRVGRSMAPIRASASSEGEIVLEQPKKD 60

DB 1 MATVRIPTISCIHHTFRSOSPRTFARIRVGRSMAPIRASASSEGEIVLEQPKKD 60

QY 61 DKKLQKGIAEFYDESSGLMENIGDHHHGFYDSSTVSLSDHRAAQRMIQESLRPAS 120

DB 61 DKKLQKGIAEFYDESSGLMENIGDHHHGFYDSSTVSLSDHRAAQRMIQESLRPAS 120

QY 121 VSEERSKPKSTIVDVGCGIGSSRYLAKKPGATSVGITLSPVQORANALAAOGLADKV 180

DB 121 VSEERSKPKSTIVDVGCGIGSSRYLAKKPGATSVGITLSPVQORANALAAOGLADKV 180

QY 181 SFQVADALQOPFSDGQFDLVMSMESGEHMPDKAKFVGEIARVAAPGAIITIVTCHRDIG 240

DB 181 SFQVADALQOPFSDGQFDLVMSMESGEHMPDKAKFVGEIARVAAPGAIITIVTCHRDIG 240

QY 241 PDEOSLHPWEDDLKKIDATYLLPAMCSTSDYVKLLQSLQDISEMSRFPVAFMPAV 300

DB 241 PDEOSLHPWEDDLKKIDATYLLPAMCSTSDYVKLLQSLQDISEMSRFPVAFMPAV 300

QY 301 IRSATFWKGLSSLSGGKTIKALAMPMLTEGKKDLIKFAITTCRPE 350

DB 301 IRSATFWKGLSSLSGGKTIKALAMPMLTEGKKDLIKFAITTCRPE 350

RESULT 2

AAY96462

ID AAY96462 standard; Protein: 349 AA.

XX AAY96462;

AC

XX 12-SEP-2000 (first entry)

DT

XX Soybean gamma tocopherol methyltransferase.

DE

XX Vitamin E: alpha-tocopherol; biosynthesis; enzyme; inhibitor; herbicide;

KW gamma-tocopherol methyltransferase.

KM

XX Glycine max.

OS

XX WO200032757-A2.

PN

XX 08-JUN-2000.

PD

XX 02-DEC-1999; 99WO-US28588.

PF

XX 03-DEC-1998; 98US-0110781.

PR

XX (DUPO) DU PONT DE NEMOURS & CO E I.

PA

XX Cahoon RE, Coughlan SJ, Miao G, Rafalski JA;

PI

DR WPI: 2000-412309/35.

DR N-PSDB: AAA29154.

XX

PT Polynucleotide encoding a vitamin E biosynthetic enzyme polypeptide

PT such as gamma tocopherol methyltransferase, useful for synthesizing

PT vitamin E or as a target for design and discovery of herbicides

XX

PS Claim 11; Page 52-53; 82pp; English.

PS

XX AAY96459-77 show vitamin E (alpha-tocopherol) biosynthetic enzymes or

CC their fragments. The cDNA clones were identified by BLAST searches based

CC on similarity to Synecocystis sp. and Arabidopsis thaliana vitamin E

CC biosynthetic enzymes. The enzymes are useful for synthesizing

CC plastoquinones or tocopherols (especially vitamin E). Vitamin E is

CC required for plant growth, therefore the enzymes may be used for the

CC discovery of new herbicides. The enzymes can be used in methods to

CC evaluate potential inhibitors, which may have use as herbicides.

CC Additionally, transgenic expression of, e.g. gamma-tocopherol

CC methyltransferase affords the ability to manipulate tocopherol levels as

CC desired for a particular application.

CC

XX Sequence 349 AA;

SO

Query Match 91.2%; Score 1668.5; DB 21; Length 349;

Best Local Similarity 93.1%; Pred. No. 1.9e-156;

Matches 323; Conservative 5; Mismatches 10; Indels 9; Gaps 1;

QY 1 MATVRIPTISCIHHTFRSOSPRTFARIRVGRSMAPIRASASSEGEIVLEQPKKD 60

DB 1 MATVRIPTISCIHHTFRSOSPRTFARIRVGRSMAPIRASASSEGEIVLEQPKKD 60

QY 61 DKKLQKGIAEFYDESSGLMENIGDHHHGFYDSSTVSLSDHRAAQRMIQESLRPAS 120

DB 61 DKKLQKGIAEFYDESSGLMENIGDHHHGFYDSSTVSLSDHRAAQRMIQESLRPAS 120

QY 121 VSEERSKPKSTIVDVGCGIGSSRYLAKKPGATSVGITLSPVQORANALAAOGLADKV 180

DB 121 VSEERSKPKSTIVDVGCGIGSSRYLAKKPGATSVGITLSPVQORANALAAOGLADKV 180

QY 241 PDEOSLHPWEDDLKKIDATYLLPAMCSTSDYVKLLQSLQDISEMSRFPVAFMPAV 300

DB 241 PDEOSLHPWEDDLKKIDATYLLPAMCSTSDYVKLLQSLQDISEMSRFPVAFMPAV 300

QY 301 IRSATFWKGLSSLSGGKTIKALAMPMLTEGKKDLIKFAITTCR 347

DB 301 IRSATFWKGLSSLSGGKTIKALAMPMLTEGKKDLIKFAITTCR 338

RESULT 3

AAY96473

ID AAY96473 standard; Protein: 366 AA.

XX AAY96473;

AC

XX 12-SEP-2000 (first entry)

DT

XX Soybean gamma-tocopherol methyltransferase.

DE

XX Vitamin E: alpha-tocopherol; biosynthesis; enzyme; inhibitor; herbicide;

KW gamma-tocopherol methyltransferase.

KM

XX Triticum aestivum.

OS

XX Key Location/Qualifiers

FT Misc-difference 5

FT Misc-difference 5

FT Misc-difference 45

FT /note- "encoded by GNC"

FT /note- "encoded by CNG"

XX

PN WO200032757-A2.
 XX 08-JUN-2000.
 PD 02-DEC-1999; 99WO-US28588.
 XX 03-DEC-1998; 98US-0110781.
 XX (DUP0) DU PONT DE NEMOURS & CO E I.
 PA Cahoon RE, Coughlan SJ, Miao G, Rafalski JA;
 PI WPI: 2000-412309/35.
 DR N-PSDB: AAA29165.
 XX Polynucleotide encoding a vitamin E biosynthetic enzyme polypeptide
 PT such as gamma tocopherol methyltransferase, useful for synthesizing
 PT vitamin E or as a target for design and discovery of herbicides
 XX
 PS Claim 11; Page 67-68; 82pp; English.
 XX
 CC AA96459-77 encode vitamin E (alpha-tocopherol) biosynthetic enzymes or
 CC their fragments. The cDNA clones were identified by BLAST searches based
 CC on similarity to Synecocystis sp. and Arabidopsis thaliana vitamin E
 CC biosynthetic enzymes. The enzymes are useful for synthesizing
 CC plastoquinones or tocopherols (especially vitamin E). Vitamin E is
 CC required for plant growth, therefore the enzymes may be used for the
 CC discovery of new herbicides. The enzymes can be used in methods to
 CC evaluate potential inhibitors, which may have use as herbicides.
 CC Additionally, transgenic expression of, e.g. gamma-tocopherol
 CC methyltransferase affords the ability to manipulate tocopherol levels as
 CC desired for a particular application.
 CC
 XX
 SQ Sequence 366 AA:

Query Match 64.1%; Score 1173; DB 21; Length 366;
 Best Local Similarity 70.4%; Pred. No. 2.2e-107;
 Matches 224; Conservative 33; Mismatches 61; Indels 0; Gaps 0;

QY 33 PPSMAPIRASASSEGEIVLEOKPKDKKKIQAETDESSGLWENIMCDHMHNGF 92
 DB 34 PPSVPLPLPARXLRAPDGVVDRPGDPAAPPGLEKAGLYDESSGIMESIMCEHMHNGF 93
 QY 93 YOSDSVSLSDHRAQIRMOESLRFASVSEERSKMPKSIYDVGCGISSRYLAKKFGA 152
 DB 94 YOSGEASMSDHRRAQIRMEELAFRAVDDPTNKEKITVDVCGGSSRYLANKYGA 153
 QY 153 TSVGITLSPVOAQRANALAAAGIADKVSFOVADALQPPSDQFDLWMSSEGEHMPDK 212
 DB 154 QCSGITLSPVOAQRANALAAAGIADKVSFOVADALQPPSDQFDLWMSSEGEHMPDK 213
 QY 213 AKFVGLARVAARFAGATIIITWCHRDGLPDEGSLHWEODLKKICDAYYLPAWCSTSDY 272
 DB 214 OKFVSELARVAARFAGATIIITWCHRNLPASEDSLKPDELMLKKICDAYYLPWCSPSDY 273
 QY 273 VKILOSLLODIKSEDMRSRVAPFVAVTASATFWKGLSSLLSGOKTIGALAMPMTIE 332
 DB 274 VKIAESLSLEDITADMSSEVAPFVAVTASATFWKGLSSLLSGOKTIGALAMPMTIE 333
 QY 333 GYKGLIKRSIITCRKRPQ 351
 DB 334 GYKGLIKRSIITCRKRPQ 351

RESULT 4
 AA96469
 ID AA96469 standard; Protein; 352 AA.
 AC AA96469;
 XX
 XX 12-SEP-2000 (first entry)
 XX
 DE Corn gamma-tocopherol methyltransferase.

XX
 KW Vitamin E: alpha-tocopherol; biosynthesis; enzyme; inhibitor; herbicide;
 KV gamma-tocopherol methyltransferase.
 XX
 OS Zea mays.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 152
 FT /note= "encoded by NCBI"
 XX
 PN WO200032757-A2.
 XX 08-JUN-2000.
 PD 02-DEC-1999; 99WO-US28588.
 XX 03-DEC-1998; 98US-0110781.
 XX (DUP0) DU PONT DE NEMOURS & CO E I.
 PA Cahoon RE, Coughlan SJ, Miao G, Rafalski JA;
 PI WPI: 2000-412309/35.
 DR N-PSDB: AAA29161.
 XX Polynucleotide encoding a vitamin E biosynthetic enzyme polypeptide
 PT such as gamma tocopherol methyltransferase, useful for synthesizing
 PT vitamin E or as a target for design and discovery of herbicides
 XX
 PS Claim 11; Page 60-62; 82pp; English.
 XX
 CC AA96459-77 show vitamin E (alpha-tocopherol) biosynthetic enzymes or
 CC their fragments. The cDNA clones were identified by BLAST searches based
 CC on similarity to Synecocystis sp. and Arabidopsis thaliana vitamin E
 CC biosynthetic enzymes. The enzymes are useful for synthesizing
 CC plastoquinones or tocopherols (especially vitamin E). Vitamin E is
 CC required for plant growth, therefore the enzymes may be used for the
 CC discovery of new herbicides. The enzymes can be used in methods to
 CC evaluate potential inhibitors, which may have use as herbicides.
 CC Additionally, transgenic expression of, e.g. gamma-tocopherol
 CC methyltransferase affords the ability to manipulate tocopherol levels as
 CC desired for a particular application.
 CC
 XX
 SQ Sequence 352 AA:

Query Match 63.5%; Score 1162; DB 21; Length 352;
 Best Local Similarity 66.2%; Pred. No. 2.6e-106;
 Matches 225; Conservative 33; Mismatches 72; Indels 10; Gaps 2

QY 20 SOSPPFARIRVGPSPMAPIRASASSEGEIVLEOKPKDKKK-----LQKIAE 71
 DB 10 SOSRSCLACRKSHTRAVSHPVRSRRRLRAVAVSLRPMASSTQAQAPATAPPLKEGIC 69
 QY 72 FYDESSGLWENIMCDHMHNGFVDSVSLSDHRAQIRMOESLRFASV--SEERSKMP 129
 DB 70 LYDESSGLWENIMCDHMHNGFVDSVSEASADHRRRAQIRMEELAFAGVAPASDDPEKTP 129
 QY 130 KSTIVDVGCGIGSSRYLAKKFGATSVGITLSPVOAQRANALAAAGIADKVSFOVADALQ 169
 DB 130 KTIYDVGCGIGSSRYLAKKFGATSVGITLSPVOAQRANALAAAGIADKVSFOVADALQ 169
 QY 190 QPESDQFDLWMSSEGEHMPDKAKFVGLARVAARFAGATIIITWCHRDGLPDEGSLHWP 249
 DB 190 QPESDQFDLWMSSEGEHMPDKAKFVGLARVAARFAGATIIITWCHRNLPASEDSLKP 249
 QY 250 EODLKKICDAYYLPAWCSTSDYVKILOSLLODIKSEDMRSRVAPFVAVTASATFWKGL 309
 DB 250 ELSLARICDAYYLPWCSPSDYVNTAKSLSDITADMSSEVAPFVAVTASATFWKGL 309
 QY 310 LSSLLSGOKTIGALAMPMTIEGYKGLIKRSIITCRKRP 349
 DB 310 FTSLLTGWKTIGAMVPMILOGYKGLIKRSIITCRKRP 349

RESULT 5
AAG51381
ID AAG51381 standard; Protein: 348 AA.
XX AAG51381;
AC AAG51381;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 65206.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hydridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-012180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145911.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147411.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.

```

PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156559.
PR 26-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159328.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 63.1%; Score 1154.5; DB 21; Length 348;
 Best Local Similarity 69.4%; Pred. No. 1.4e-105;
 Matches 227; Conservative 28; Mismatches 61; Indels 11; Gaps 4;

```

OY 32 GPRNAPIRASASSE-----RGEIVLEOKRKDDKKLKGIAEFYDESSGLWENING 85
DB 23 GSKSLLFRSPSSSSSVSMTRGNVAVAAATSTE--ALRKGIAEFYNETSGLWEEING 80
OY 86 DHMHGFTDSDVSLSD--HRAAQIRMIQESLRPASVS--EERSKMPKSIYVGGGIGGS 142
DB 81 DHMHGFTDSDVSLSDSGHKEAQIRMTESLRFAGVDEEERKIKRVVDVGGGIGGS 140
OY 143 SHYLAKKFGATSVGITLSPVQARANALAAAGLADKVSFOVADALQDFPSDGFPLVMS 202
DB 141 SHYLAKKFGAECIGITLSPVQARANDLAAAGSLAKKASFOVADALQDFPSDGFPLVMS 200

```

```

OY 203 MESGEHMPDKAKFVCELARVAAPGAIITVTCRDLCPDQSLHPWEDLLKTCIDAVY 262
DB 201 MESGEHMPDKAKFVCELARVAAPGRIITVTCRHNLSAGEALQPMQONITLDRICTFY 260
OY 263 LPANCSTSDYKLLQSLSDQIKSEDSRFAPMPAPVIRAPFVWKGSLLSSGOKTIK 322
DB 261 LPANCSTDDYVNLQSHSLQDKADWSENVAFPMPAVIRATLWKGLVLSRSGMSIK 320
OY 323 GALAMPMEIGYKKDLIKFAITTCRKP 349
DB 321 GALTPLMTEGKKGKVGKIGITTCOKP 347

```

RESULT 6

ABB91452 standard; protein: 348 AA.

ABB91452;

31-MAY-2002 (first entry)

Herbicideally active polypeptide SEQ ID NO 663.

Herbicideal; plant; agriculture; herbicide.

Arabidopsis thaliana.

WO200210210-A2.

07-FEB-2002.

28-AUG-2001; 2001WO-EP09892.

28-AUG-2001; 2001WO-EP09892.

(FARB) BAYER AG.

Tietjen K, Weidner M;

WPI; 2002-269010/31.

Identifying plant target proteins for herbicideally active compounds, comprising aligning and comparing nucleic acid or amino acid sequences from plant with nucleic acid or amino acid sequences from non-plant organisms -

Claim 5; SEQ ID NO 663; 261pp + Sequence Listing; English.

The invention relates to identifying target proteins

(ABB90790-ABB94016) for herbicideally active compounds, comprising aligning and comparing nucleic acid or amino acid sequences from plant with nucleic acid or amino acid sequences from non-plant organisms using suitable search parameters, where plant sequences having an E-value greater by a factor of 3 than the E-value of most similar non-plant sequences are selected. The polypeptides or nucleic acids encoding them are useful for identifying modulators. The identified modulators are useful as herbicides.

Sequence 348 AA;

Query Match 63.1%; Score 1154.5; DB 23; Length 348;

Best Local Similarity 69.4%; Pred. No. 1.4e-105; Indels 11; Gaps 4

Matches 227; Conservative 28; Mismatches 61; Indels 11; Gaps 4

```

OY 32 GPRNAPIRASASSE-----RGEIVLEOKRKDDKKLKGIAEFYDESSGLWENING 85
DB 23 GSKSLLFRSPSSSSSVSMTRGNVAVAAATSTE--ALRKGIAEFYNETSGLWEEING 80
OY 86 DHMHGFTDSDVSLSD--HRAAQIRMIQESLRPASVS--EERSKMPKSIYVGGGIGGS 142
DB 81 DHMHGFTDSDVSLSDSGHKEAQIRMTESLRFAGVDEEERKIKRVVDVGGGIGGS 140

```


PR	04-AUG-1999;	9905-0147302;
PR	05-AUG-1999;	9905-0147192;
PR	05-AUG-1999;	9905-0147260;
PR	06-AUG-1999;	9905-0147303;
PR	06-AUG-1999;	9905-0147416;
PR	09-AUG-1999;	9905-0147493;
PR	09-AUG-1999;	9905-0147935;
PR	10-AUG-1999;	9905-0148171;
PR	11-AUG-1999;	9905-0148319;
PR	12-AUG-1999;	9905-0148341;
PR	13-AUG-1999;	9905-0148565;
PR	13-AUG-1999;	9905-0148684;
PR	16-AUG-1999;	9905-0149368;
PR	17-AUG-1999;	9905-0149175;
PR	18-AUG-1999;	9905-0149426;
PR	20-AUG-1999;	9905-0149722;
PR	20-AUG-1999;	9905-0149723;
PR	20-AUG-1999;	9905-0149929;
PR	23-AUG-1999;	9905-0149902;
PR	23-AUG-1999;	9905-0149930;
PR	25-AUG-1999;	9905-0150566;
PR	26-AUG-1999;	9905-0150584;
PR	27-AUG-1999;	9905-0151065;
PR	27-AUG-1999;	9905-0151065;
PR	27-AUG-1999;	9905-0151080;
PR	31-AUG-1999;	9905-0151303;
PR	01-SEP-1999;	9905-0151930;
PR	07-SEP-1999;	9905-0151363;
PR	10-SEP-1999;	9905-0153070;
PR	13-SEP-1999;	9905-0153758;
PR	15-SEP-1999;	9905-0154018;
PR	16-SEP-1999;	9905-0154039;
PR	20-SEP-1999;	9905-0154779;
PR	22-SEP-1999;	9905-0155133;
PR	23-SEP-1999;	9905-0155486;
PR	24-SEP-1999;	9905-0155569;
PR	28-SEP-1999;	9905-0156548;
PR	29-SEP-1999;	9905-0156596;
PR	04-OCT-1999;	9905-0157117;
PR	05-OCT-1999;	9905-0157753;
PR	06-OCT-1999;	9905-0157863;
PR	07-OCT-1999;	9905-0158029;
PR	08-OCT-1999;	9905-0158232;
PR	12-OCT-1999;	9905-0158369;
PR	13-OCT-1999;	9905-0159293;
PR	13-OCT-1999;	9905-0159294;
PR	14-OCT-1999;	9905-0159329;
PR	14-OCT-1999;	9905-0159329;
PR	14-OCT-1999;	9905-0159331;
PR	14-OCT-1999;	9905-0159637;
PR	14-OCT-1999;	9905-0159638;
PR	18-OCT-1999;	9905-0159584;
PR	21-OCT-1999;	9905-0160741;
PR	21-OCT-1999;	9905-0160767;
PR	21-OCT-1999;	9905-0160768;
PR	21-OCT-1999;	9905-0160770;
PR	21-OCT-1999;	9905-0161361;
PR	26-OCT-1999;	9905-0161361;
PR	28-OCT-1999;	9905-0161993;
PR	29-OCT-1999;	9905-0162142;

Query Match	62.9%	Score 1151.5	DB 21	Length 308				
Best Local Similarity	71.3%	Pred. No. 2.3e-105						
Matches 224	Conservative 24	Mismatches 47	Indels 19	Gaps				
OY	IRAAASASERCEIYLEQPKDKKKLQKQIAEFYDESSGLMENIWMGHMHGFYDSDT 98							
DB	10 VAAATSTE-----ALRKQIAEFYMETSGIMEIWMGHMHGFYDPS 53							
OY	99 VLSLD--HRAAQIRMIQESLRPASVS-EERSKWPKSIYDVCGIGGSSRYLAKKFGATSV 155							
DB	54 VQLSDSGKHEAQIMIESERLRFAGVTDDEEEKKIKKKVVYGVCGIGGSSRYLAKSGAGECI 113							
OY	156 GTTSLPQAOANMLAAQGLADLAKVSQVADALQGFSDGPDIDVWSMSSGSHMPDKAF 215							
DB	114 GTTSLPQAKRANDLAAQGLADLAKVSQVADALQGFSDGPDIDVWSMSSGSHMPDKAF 173							
OY	216 VGLLARVAAPCAIIITITWCHRDLPDGEOSLHPEODLLKIKCDAYVLPAMCSTSDYVKL 275							
DB	174 VKELRVAAAPCGRIITITWCHRNLSAGEALQPMQNIIDKIKCTFYLPAWCSTDYVNL 233							
OY	276 LQSLSLQDIKSEDSRFVAPRPVYIRSAFTWKGSLSSGOKTIGALAMPIMEGK 335							
DB	234 LQSHSLQDIKCADSENVAPFPVAVIRTAITLWKGVLSSLRSGKMSIKGALTPLMIEGK 293							
OY	336 KDLKFAITTCRK 349							
DB	294 KGVIRFGITTCRK 307							
RESULT 8								
AAW95017								
ID	AAW95017 standard; Protein: 348 AA.							
AC	AAW95017;							
XX	21-MAY-1999 (first entry)							
DE	Arabidopsis gamma-tocopherol methyltransferase (gamma-TMT).							
XX	Gamma-tocopherol methyltransferase; gamma-TMT; SLR0089, 165H577; meat-							
KW	transgenic plant; alpha-tocopherol; gamma-tocopherol; forage plant.							
OS	Arabidopsis sp.							
XX	W09904622-A1.							
PN	04-FEB-1999.							
XX	22-JUL-1998; 98WO-US15137.							
XX	17-JUL-1998; 98US-0053819.							
PR	25-JUL-1997; 97US-0053819.							
PR	26-JAN-1998; 98US-0072497.							
XX	(UYNE-) UNIV NEVADA.							
XX	Dellapenna D, Shintani DK;							
XX	WPI; 1999-142456/12.							
DR	N-PSDB; AAX17789.							
XX	Newly isolated DNA fragment comprising a ^c-tocopherol (vitamin E)							
PT	methyltransferase coding sequence - useful for producing							
PT	^a-tocopherol, and transgenic plants, seeds and oils with an altered							
PT	tocopherol profile							
XX	Example 5; Page 36-38; 46pp; English.							
PS	The invention provides DNA sequences encoding gamma-tocopherol							
XX	methyltransferases (gamma-TMT). The gamma-TMT encoding genes SLR0089							
CC	and 165H577 are isolated from <i>Synechocystis</i> and <i>Arabidopsis</i> species							
CC	respectively. The DNA fragments are useful for producing transgenic							

CC plants containing a higher alpha-tocopherol:gamma-tocopherol ratio. They
 CC are also useful for producing alpha-tocopherol and plants with increased
 CC gamma-tocopherol, which may be useful in certain industries such as the
 CC meat industry e.g. for developing forage plants to feed animals. The
 CC production of transgenic plants (and seeds) with a higher alpha-
 CC tocopherol:gamma-tocopherol ratio, using the gamma TMTs, will increase
 CC the level of alpha-tocopherol in the human diet, in addition to enhancing
 CC the stability and shelf life of plants and plant products. Increased
 CC levels of alpha-tocopherol will also increase meat quality and extend
 CC shelf life of post-processed meat products. Plants with a higher alpha-
 CC tocopherol:gamma-tocopherol ratio may also produce advantageous
 CC phenotypes. The present sequence represents the Arabidopsis gamma-TMT.
 XX

SO Sequence 348 AA;

Query Match 62.9%; Score 1151.5; DB 20; Length 348;
 Best Local Similarity 69.1%; Pred. No. 2.8e-105;
 Matches 226; Conservative 29; Mismatches 61; Indels 11; Gaps 4;

OY 32 GPRSMAPIRASASSE-----RGEIVLEQPKKDDKKKLOKGAIEFYDESSGIMENIWG 85
 DB 23 GSKSLLFRSPSSSSSVSMTTTNGNVAANAATSTE--ALRKGIAEFYNETSGIMEIING 80
 OY 86 DHMHGFIYDSDSTVSLSD--HRAQIRMIOESLRFASVS--EERSKMPKSIYDVGCGIGS 142
 DB 81 DHMHGFIYDSDSTVSLSDSGHKEAOIRNIESLRFAGVTDEEEKIKKIVDVCGLIGS 140
 OY 143 SRYIAKFGATSVGITTSPVOAORANLAAAGLADRVSFQVADALQOPFSDGQFDLYWS 202
 DB 141 SRYIAKFGAECIGITLSPVOAKRANLAAOSLSHKASQVADALQOPFSDGQFDLYWS 200
 OY 203 MESSEHMPDKAKFVGEIARVAARCAIITIVWCHRDLPDQSLHPMEDOLKIKICAYY 262
 DB 201 MESSEHMPDKAKFVGEIARVAARCAIITIVWCHRNLSAGEALQPEONLIDKICTFY 260
 OY 263 LPACSTSDYVKLQSLSDIKSEDSRFVAPFPAVIRSAFTWKGSLSSGOKTYK 322
 DB 261 LPACSTSDYVKLQSLSDIKSEDSRFVAPFPAVIRSAFTWKGSLSSGOKTYK 320
 OY 323 GALAMPIMEGKKRDLIKFAITTCRKP 349
 DB 321 GALTMPIMEGKKGVKIFGITTCOKP 347

RESULT 9

AAV77937 ID AAV77937 standard; protein; 348 AA.

XX AAV77937;

DT 14-JUN-2000 (first entry)

XX A. thaliana environmental stress tolerance related protein.

KM Environmental stress; plant; transgenic plant; anaerobic; flooding; cold;
 KM dehydration; drought; heat stress; salinity; osmotolerance.

XX Arabidopsis thaliana.

XX WO200008187-A2.

PD 17-FEB-2000.

XX 04-AUG-1999; 99WO-EP05652.

PR 04-AUG-1998; 98EP-0202634.

PA (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

XX Lee JH, Verbruggen N;

XX WPI; 2000-205726/18.

DR N-PSDB; AA298317.

XX Isolation of polynucleic acids useful for producing transgenic plant by
 PT isolating genes involved in tolerance to environmental stress
 XX Claim 12; Page 130-132; 312pp; English.

CC The invention relates to isolation of coding sequences and/or genes
 CC involved in tolerance to environmental stress in plants. The sequences
 CC (AA298305-298365) are useful for producing a transgenic plant having
 CC enhanced tolerance or resistance to environmental stress conditions such
 CC as anaerobic, flooding, cold, dehydration, drought, heat stress or
 CC salinity. This is useful for producing improved yield, growth,
 CC development and productivity under environmental stress conditions, and
 CC also provides growth of crops in areas where they cannot grow without
 CC the induced osmotolerance. Sequences AAV77925-984 represent polypeptide
 CC sequences from A. thaliana that are encoded by the genes involved in
 CC environmental stress tolerance.
 XX

SO Sequence 348 AA;

Query Match 62.9%; Score 1151.5; DB 21; Length 348;
 Best Local Similarity 69.1%; Pred. No. 2.8e-105;
 Matches 226; Conservative 29; Mismatches 61; Indels 11; Gaps 4;

OY 32 GPRSMAPIRASASSE-----RGEIVLEQPKKDDKKKLOKGAIEFYDESSGIMENIWG 85
 DB 23 GSKSLLFRSPSSSSSVSMTTTNGNVAANAATSTE--ALRKGIAEFYNETSGIMEIING 80
 OY 86 DHMHGFIYDSDSTVSLSD--HRAQIRMIOESLRFASVS--EERSKMPKSIYDVGCGIGS 142
 DB 81 DHMHGFIYDSDSTVSLSDSGHKEAOIRNIESLRFAGVTDEEEKIKKIVDVCGLIGS 140
 OY 143 SRYIAKFGATSVGITTSPVOAORANLAAAGLADRVSFQVADALQOPFSDGQFDLYWS 202
 DB 141 SRYIAKFGAECIGITLSPVOAKRANLAAOSLSHKASQVADALQOPFSDGQFDLYWS 200
 OY 203 MESSEHMPDKAKFVGEIARVAARCAIITIVWCHRDLPDQSLHPMEDOLKIKICAYY 262
 DB 201 MESSEHMPDKAKFVGEIARVAARCAIITIVWCHRNLSAGEALQPEONLIDKICTFY 260
 OY 263 LPACSTSDYVKLQSLSDIKSEDSRFVAPFPAVIRSAFTWKGSLSSGOKTYK 322
 DB 261 LPACSTSDYVKLQSLSDIKSEDSRFVAPFPAVIRSAFTWKGSLSSGOKTYK 320
 OY 323 GALAMPIMEGKKRDLIKFAITTCRKP 349
 DB 321 GALTMPIMEGKKGVKIFGITTCOKP 347

RESULT 10

AAO17426 ID AAO17426 standard; Protein; 348 AA.

XX AAO17426;

DT 11-JUL-2002 (first entry)

XX A thaliana gamma-tocopherol methyltransferase.

XX Vitamin E; homogenisate; HG; homogenisate-1,2-dioxygenase; HGD;
 KM methylacetacetate-isomerase; MAAI; fumarylacetacetate-hydrolase; FAAH;
 KM transgenic; nutrition; antioxidant; cardiovascular disease; cancer;
 KM immune system; generalised age-related degeneration; animal feed;
 KM meat quality; cosmetics; growth regulator; herbicide; cardiant;
 KM cytosolic; immunostimulant; enzyme.

XX Arabidopsis thaliana.

XX WO200231173-A2.

XX 18-APR-2002.

PF 18-SEP-2001; 2001WO-EP10779.

RESULT 12
AAG07857 standard; Protein: 308 AA.
ID AAG07857
XX AAG07857;
AC AAG07857;
XX 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 5169.
XX
XX Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125768.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135623.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148563.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.

```
PR 20-AUG-1999; 99US-0149722..
PR 20-AUG-1999; 99US-0149723..
PR 20-AUG-1999; 99US-0149929..
PR 23-AUG-1999; 99US-0149902..
PR 23-AUG-1999; 99US-0149930..
PR 25-AUG-1999; 99US-0150566..
PR 26-AUG-1999; 99US-0150884..
PR 27-AUG-1999; 99US-0151065..
PR 27-AUG-1999; 99US-0151066..
PR 27-AUG-1999; 99US-0151080..
PR 30-AUG-1999; 99US-0151303..
PR 31-AUG-1999; 99US-0151438..
PR 01-SEP-1999; 99US-0151930..
PR 07-SEP-1999; 99US-0152363..
PR 10-SEP-1999; 99US-0153070..
PR 13-SEP-1999; 99US-0153758..
PR 15-SEP-1999; 99US-0154018..
PR 16-SEP-1999; 99US-0154039..
PR 20-SEP-1999; 99US-0154779..
PR 22-SEP-1999; 99US-0155139..
PR 23-SEP-1999; 99US-0155486..
PR 24-SEP-1999; 99US-0155659..
PR 28-SEP-1999; 99US-0156458..
PR 29-SEP-1999; 99US-0156596..
PR 04-OCT-1999; 99US-0157117..
PR 05-OCT-1999; 99US-0157753..
PR 06-OCT-1999; 99US-0157865..
PR 07-OCT-1999; 99US-0158029..
PR 08-OCT-1999; 99US-0158232..
PR 12-OCT-1999; 99US-0158369..
PR 13-OCT-1999; 99US-0159293..
PR 13-OCT-1999; 99US-0159294..
PR 13-OCT-1999; 99US-0159295..
PR 14-OCT-1999; 99US-0159329..
PR 14-OCT-1999; 99US-0159330..
PR 14-OCT-1999; 99US-0159331..
PR 14-OCT-1999; 99US-0159637..
PR 14-OCT-1999; 99US-0159638..
PR 18-OCT-1999; 99US-0159584..
PR 21-OCT-1999; 99US-0160741..
PR 21-OCT-1999; 99US-0160767..
PR 21-OCT-1999; 99US-0160768..
PR 21-OCT-1999; 99US-0160770..
PR 21-OCT-1999; 99US-0160814..
PR 21-OCT-1999; 99US-0160815..
PR 22-OCT-1999; 99US-0160980..
PR 22-OCT-1999; 99US-0160981..
PR 22-OCT-1999; 99US-0160989..
PR 25-OCT-1999; 99US-0161404..
PR 25-OCT-1999; 99US-0161405..
PR 25-OCT-1999; 99US-0161406..
PR 26-OCT-1999; 99US-0161359..
PR 26-OCT-1999; 99US-0161360..
PR 26-OCT-1999; 99US-0161361..
PR 28-OCT-1999; 99US-0161920..
PR 28-OCT-1999; 99US-0161921..
PR 28-OCT-1999; 99US-0161993..
PR 29-OCT-1999; 99US-0162142..

Query Match      62.5%; Score 1144.5; DB 21; Length 308;
Best Local Similarity 70.7%; Pred. No. 1,2e-104;
Matches 222; Conservative 26; Mismatches 47; Indels 19; Gaps 3;
```

```
DB 114 GTTSPVOAKRANDLAAASLSHKASFOYADALDKRFEDGKTFDLYWSMESGEMHPDKAF 173
OY 216 VGEIARVAAPAGAIITIIIVTWCHRDLPDEOSLHPBEDOLKICIDAVYIIPAWCSTSDYVL 275
DB 174 VKELRVAAAPGRIITIIIVTWCHRNLSAGEEALQPMEDNIIIDKICTFYIPLAWCSTDYVNL 233
OY 276 LOSLSLDIDKSEDMRFPAPFPVAVIRSAFTWKGLSSLLSGOKITKALAMPIMIEGYK 335
DB 234 LOSLSLDIDKICADMESENAFPAPVAVIRTAFTWKGLVSLRSRGSIKGALTPLMIEGYK 293
OY 336 KDLIKFAIITCRKP 349
DB 294 KGVIKFGIITCRKP 307

RESULT 13
AAG51383
ID AAG51383 standard; Protein: 266 AA.
XX AAG51383;
AC
XX
DF 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 65208.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
PD
XX 06-SEP-2000.
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825..
PR 05-MAR-1999; 99US-0123180..
PR 09-MAR-1999; 99US-0123548..
PR 23-MAR-1999; 99US-0125788..
PR 25-MAR-1999; 99US-0126264..
PR 29-MAR-1999; 99US-0126785..
PR 01-APR-1999; 99US-0127462..
PR 06-APR-1999; 99US-0128234..
PR 16-APR-1999; 99US-0128714..
PR 19-APR-1999; 99US-0129845..
PR 21-APR-1999; 99US-0130077..
PR 21-APR-1999; 99US-0130449..
PR 23-APR-1999; 99US-0130510..
PR 23-APR-1999; 99US-0130891..
PR 28-APR-1999; 99US-0131449..
PR 30-APR-1999; 99US-0132048..
PR 04-MAY-1999; 99US-0132407..
PR 04-MAY-1999; 99US-0132484..
PR 05-MAY-1999; 99US-0132485..
PR 06-MAY-1999; 99US-0132486..
PR 06-MAY-1999; 99US-0132487..
PR 07-MAY-1999; 99US-0132863..
PR 11-MAY-1999; 99US-0134256..
PR 14-MAY-1999; 99US-0134218..
PR 14-MAY-1999; 99US-0134219..
PR 14-MAY-1999; 99US-0134221..
PR 14-MAY-1999; 99US-0134370..
PR 18-MAY-1999; 99US-0134376..
PR 19-MAY-1999; 99US-0134941..
PR 20-MAY-1999; 99US-0135124..
PR 21-MAY-1999; 99US-0135353..
PR 24-MAY-1999; 99US-0135629..
PR 25-MAY-1999; 99US-0136021..
PR 27-MAY-1999; 99US-0136392..
PR 28-MAY-1999; 99US-0136782..
PR 01-JUN-1999; 99US-0137222..
```

PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160817.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 29-OCT-1999; 99US-0162142.

Query Match

56.6%; Score 1035.5; DB 21; Length 266;

Best Local Similarity 75.8%; Pred. No. 5,7e-94;
Matches 201; Conservative 18; Mismatches 43; Indels 3; Gaps 2;

```
OY 88 MHGFDSDSTVSLD--HRAQIRMIQESLRASVS-EERSKPKSIIVDGGIGSSR 144
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1 MHGFDSDSDSTVSLDSDGHEKAQIRMIESLRFAGVDEDEEKKIKRVVDVGGIGSSR 60
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
OY 145 YLAKKFEATSVGTLTSVQAQRANALAAAGLADKVSFOYADALQOFSSGQDVLVSM 204
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 61 YLAKKFEATSVGTLTSVQAQRANALAAAGLADKVSFOYADALQOFSSGQDVLVSM 120
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
OY 205 SCHEMPDKAFVCELAFAVAPGAIITVTCNRDLPDEQSLHPMODLTKICDAYILP 264
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 121 SCHEMPDKAFVCELAFAVAPGAIITVTCNRDLPDEQSLHPMODLTKICDAYILP 180
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
OY 265 AKCSTDYVKLOSLSLQDIKSEDSRFVAPFVPVIRSAFTWKGLSSLSGQKTIKA 324
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 181 AKCSTDYVKLOSLSLQDIKSEDSRFVAPFVPVIRSAFTWKGLSSLSGQKTIKA 240
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
OY 325 LAMPMEGYKKDLIRFAITCRP 349
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 241 LAMPMEGYKKDLIRFAITCRP 265
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 14
AAG07858
ID AAG07858 standard; Protein; 266 AA.
XX
AC AAG07858;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 5170.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0126264.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130449.
PR 23-APR-1999; 990S-0130510.
PR 23-APR-1999; 990S-0130891.
PR 28-APR-1999; 990S-0131449.
PR 30-APR-1999; 990S-0132048.
PR 30-APR-1999; 990S-0132407.
PR 04-MAY-1999; 990S-0132484.
PR 05-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 06-MAY-1999; 990S-0132487.
PR 07-MAY-1999; 990S-0132863.
PR 11-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134370.
```

```
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135328.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139492.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139763.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144684.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145918.
```

PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146388.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155569.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159684.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 56.2%; Score 1028.5; DB 21; Length 266;
Best local similarity 75.1%; Pred. No. 2.8e-93;
Matches 199; Conservative 20; Mismatches 43; Indels 3; Gaps 2.

QY 88 MHKGFYSDSTVSLSD--HRAQIRMGESLRFASVS--ERSKMKSYIVGCGIGSSR 144
||||| ||: ||| ||: |||||: ||||| ||: ||: ||||| |||||
DB 1 MHKGFYDPSVSQVSLSDSGHKEADIRNIESLRFAGVTDEEEKIKTKVYDVCIGIGSSR 60

QY 145 YLAKFGATSVGTTSPVQAGANALAAAGLADKVSFOVALAQQPSDGFPLVMSME 204
||| ||| ||: ||||| ||||| ||| ||| ||: ||||| ||: ||: ||||| |||||
DB 61 YLAKFGAECIGITLSPVQAKRANLAAQSLSHKASFQVADALQKPEDGKFDLWMSME 120

QY 205 SCEHMPDKAKFVGEGLARVAPGAIITIVWCHRDLCPPDQSLHPWEDLLKICDAPYLP 264
||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 121 SCEHMPDKAKFVKELVRAAPGCRITITVWCHRNLSAGEALQPEQNTLDKICTFYLP 180

QY 265 AMCSTSDYVKLQSLSDIDIKSEDSRFVAPFPVARSFTWKGSLSSGQRTKCA 324
||||| ||| ||| ||||| ||||| ||| ||||| ||||| ||||| ||||| |||||
DB 181 AMCSTDDYVNLQSLSDIDIKCADNSENVAPFPVATRTALWKGVLVSLRSGMSIKCA 240

QY 325 LAMPIMIEGKKDLIKFAITTCRK 349
| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 241 LTPMIMIEGKKGVKFGITTCOKP 265

RESULT 15
AAB93888
ID AAB93888 standard; Protein; 370 AA.
XX
AAB93888;
AC
XX
DT 21-SEP-2001 (first entry)
XX
XX
DE Physcomitrella patens 78_pppocl_087_e12-259rev protein.
XX
XX
KW Tocopherol and carotenoid metabolism related protein; TCMP; synthesis;
KW Physcomitrella patens; moss; algae; microorganism; fungus; plant;
KW identification; genome mapping; modulation; evolutionary study;
KW cellular production; fine chemical.
XX
XX
OS Physcomitrella patens.
XX
PN WO200144276-A2.
XX
XX
PD 21-JUN-2001.
XX
XX
PF 14-DEC-2000; 2000WO-EP12698.
XX
XX
PR 16-DEC-1999; 99US-0171121.
XX
XX
PA (BADI) BASF PLANT SCI GMBH.
XX
XX
PI Lerchl J, Renz A, Ehrhardt F, Reindl A, Clippus P, Bischoff F;
PI Frank M, Freund A, Duvenig E, Schmidt R, Reski R, Badur R;
XX
XX
DR WPI: 2001-398121/42.
DR N-PSDB: AAH44261.
XX
XX
PT Tocopherol and carotenoid metabolism related protein (TCMP), used to
PT produce fine chemicals, is isolated from mosses, algae, microorganisms,
PT fungi, plants, or their fragments -
XX
XX
PS Claim 28; Page 122; 123pp; English.
XX
CC The present invention describes isolated tocopherol and carotenoid

CC metabolism related proteins (TCMRP) (1) from mosses or algae.
CC microorganisms or fungi, plants, or its fragments. (1) can be used as
CC enzymes in the production of fine chemicals or in the metabolism of
CC tocopherols and carotenoids. (1) also assist in transmembrane transport.
CC The fine chemicals that can be produced include lipids, fatty acids,
CC vitamins, cofactors, enzymes, amino acids, and nucleotide bases.
CC Nucleotide sequences, proteins, vectors and host cells from the present
CC invention can be used: (a) to identify mosses related to Physcomitrella
CC patens; (b) in mapping genomes of mosses related to Physcomitrella
CC patens; (c) in the modulation of TCMRP activity; (d) in evolutionary
CC studies; (e) in the determination of functional TCMRP regions; (f) and
CC in the cellular production of fine chemicals. AAH44222 to AAH44262
CC encode the Physcomitrella patens TCMRP proteins given in AAB99849 to
CC AAB99889. AAH44212 to AAH44221 represent nucleotide sequence used in
CC the exemplification of the present invention.
XX
S0 Sequence 370 AA;

Query Match 55.9%; Score 1022.5; DB 22; Length 370;
Best Local Similarity 59.4%; Pred. No. 1.8e-92;
Matches 208; Conservative 40; Mismatches 81; Indels 21; Gaps 6;

OY 11 SCIHHTFRSOSPRTFARI-----RVCPRSMAPIRASASRGEIYLEQPKKDDX 62
DB 28 SCNEIRT-RSTSV-TSAOVCGILRADDEYGRG-----VKTRSLRSGGV--RAVQRT 78

OY 63 KKLQGIAEFYDESSGLWENIWDHMHGFDSDSTVSU---SDHRAQIRMIOESLRF 118
DB 79 PELYDGIAEFYDESSGVWEGIHGHYDEIVEAVDDPDHRRQIMIKESLAY 138

OY 119 ASVSEERSKMPKSIYVCGGIGSSRYLAKKFGATSVGTTSPVOAQRANALAAAGLAD 178
DB 139 AGVPSKDLKPKTIYDVCGIGSSRYLAKKFOAKVNAITLSPVOQRAVDLTAKQGLSD 198

OY 179 KVSFOVADALQPFSDGQFDLVMSMEGHEMDKAKFVGEARVAAPGAIITITWCHRD 238
DB 199 LNFQVANAALNPFODGSDLDLVMSMEGHEMDKRRKFEVGEARVAAPGRIITLVWCHRD 258

OY 239 LGPDEOSLHPWEDLLKRICDAYVLPAMCSTSDYVKLOSLSLQDIKSEDMGRFVAPFP 298
DB 259 LKPGETSLKPDQDLKICDAFYLPAMCSPSDYVSIKDLGLQDIKSGWSEYVTFPFP 318

OY 299 AVIRSAFTWKGLSLLSGQKTIKALAMPIMEGYKKDLIKFAIITCRK 348
DB 319 AVMTKALSMEGIVGLYKSGWTMKGAFAMTLMIOGYORGLIKFAAITCRK 368

Search completed: March 18, 2003, 17:37:06
Job time : 40 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 17:36:23 ; Search time 15 Seconds

(without alignments)
686,535 Million cell updates/sec

Title: US-09-857-613a-28

Sequence: 1 MATVRIPTISCIHITFRS.....IEGYKKDLIKFAITCRKPE 350

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/3A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/3B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/3A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/3B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/3C_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/3D_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	392	21.4	318	US-09-382-906A-2	Sequence 2, Appl 1
2	254	13.9	280	US-09-029-603-2	Sequence 2, Appl 1
3	239	13.1	263	US-09-041-718-5	Sequence 5, Appl 1
4	220.5	12.0	275	US-09-036-987A-7	Sequence 7, Appl 1
5	220.5	12.0	275	US-09-370-700-7	Sequence 7, Appl 1
6	212.5	11.6	361	US-09-041-718-4	Sequence 4, Appl 1
7	203.5	11.1	383	US-09-041-718-3	Sequence 3, Appl 1
8	189	10.3	376	US-09-041-718-2	Sequence 2, Appl 1
9	152.5	8.3	283	US-09-036-987A-13	Sequence 13, Appl 1
10	152.5	8.3	283	US-09-370-700-13	Sequence 13, Appl 1
11	152	8.3	282	US-08-457-245-9	Sequence 9, Appl 1
12	107.5	5.9	233	US-09-355-166-20	Sequence 20, Appl 1
13	104.5	5.7	237	US-08-576-626A-56	Sequence 56, Appl 1
14	104.5	5.7	2544	US-08-576-626A-32	Sequence 32, Appl 1
15	100.5	5.5	251	US-08-401-068-12	Sequence 12, Appl 1
16	100.5	5.5	251	US-08-846-338-12	Sequence 12, Appl 1
17	100.5	5.5	251	US-08-411-768B-3	Sequence 3, Appl 1
18	99.5	5.4	258	US-09-134-001C-3244	Sequence 3244, Appl 1
19	92.5	5.1	985	US-08-867-941-13	Sequence 13, Appl 1
20	92.5	5.1	985	US-09-074-658-13	Sequence 13, Appl 1
21	92.5	5.1	1000	US-08-867-941-12	Sequence 12, Appl 1
22	92.5	5.1	1000	US-09-074-658-12	Sequence 12, Appl 1
23	92.5	5.1	2439	US-09-074-658-11	Sequence 11, Appl 1
24	92	5.0	308	US-08-457-245-8	Sequence 8, Appl 1
25	91.5	5.0	2013	US-08-324-977-12	Sequence 12, Appl 1
26	91.5	5.0	2013	US-08-384-616-12	Sequence 12, Appl 1
27	91.5	5.0	2013	US-08-904-686A-12	Sequence 12, Appl 1

28	91.5	5.0	2013	US-09-315-850-12	Sequence 12, Appl 1
29	91.5	5.0	2201	US-08-952-981A-2	Sequence 2, Appl 1
30	91.5	5.0	2620	US-08-324-977-32	Sequence 32, Appl 1
31	91.5	5.0	2620	US-08-384-616-32	Sequence 32, Appl 1
32	91.5	5.0	2620	US-08-904-686A-32	Sequence 32, Appl 1
33	91.5	5.0	2620	US-09-315-850-32	Sequence 32, Appl 1
34	91.5	5.0	2621	US-08-324-977-36	Sequence 36, Appl 1
35	91.5	5.0	2621	US-08-384-616-36	Sequence 36, Appl 1
36	91.5	5.0	2621	US-08-904-686A-36	Sequence 36, Appl 1
37	91.5	5.0	2621	US-09-315-850-36	Sequence 36, Appl 1
38	91.5	5.0	3010	US-08-324-977-14	Sequence 14, Appl 1
39	91.5	5.0	3010	US-08-384-616-14	Sequence 14, Appl 1
40	91.5	5.0	3010	US-08-904-686A-14	Sequence 14, Appl 1
41	91.5	5.0	3010	US-09-315-850-14	Sequence 14, Appl 1
42	91.5	5.0	3010	US-08-904-686A-12	Sequence 12, Appl 1
43	91.5	5.0	3010	US-08-904-686A-14	Sequence 14, Appl 1
44	91.5	5.0	3010	US-09-315-850-12	Sequence 12, Appl 1
45	91.5	5.0	3010	US-09-315-850-14	Sequence 14, Appl 1

ALIGNMENTS

RESULT 1
US-09-382-906A-2
Sequence 2, Application US/09382906A
Patent No. 6448475

GENERAL INFORMATION:
APPLICANT: Deltapenna, Dean

APPLICANT: Shintani, David

TITLE OF INVENTION: Manipulation of Tocopherol Levels in Transgenic Plants

FILE REFERENCE: 920905, 90032

CURRENT APPLICATION NUMBER: US/09/382,906A

PRIOR FILING DATE: 1999-08-25

PRIOR APPLICATION NUMBER: 60/097,863

NUMBER OF SEQ ID NOS: 4

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 2

LENGTH: 318

TYPE: PRT

ORGANISM: Synechocystis PCC6803

US-09-382-906A-2

Query Match 21.4%; Score 392; DB 4; Length 318;

Best Local Similarity 36.1%; Pred. No. 4.3e-34; Matches 92; Conservative 40; Mismatches 87; Indels 36; Gaps 11

QY	69	IAEYDE--SSGLMENIWDHMHGFYDSVSLSDHRAAQIMIGESLRFASVSEBS	126
DB	36	VANVADWTEDEGILEYWGDIHILGHY-GDEPVAKD-----FIOSKIDFV--NAMA	83
QY	127	KW-----PKSIVDCCGIGSSRYLAKFKFATSGITLSPVQOARANLAAOGLAD	178
DB	84	QMGGLDLPPTGYVLDGCGISRLIAKQIGNVGTITLSPQOVKATELTTPPDVTA-	142
QY	179	KVSFOVADALQOPSCDFDLVWMSGESHPDRAKFEVGLARVAAPCAITITVWCHRD	238
DB	143	--KRAVDAMALSPFDSFDVWSEADPHRPAKFAKELRLVVKPGILVADMDNRD	200
QY	239	LGPEEQ-SLAFMEDDLKKIKIDAYLLPAMCSTSDYVKKLSLSD--IKEDMSRFVAP	295
DB	201	--DROVLEINFEKRPVNRKLDOWSHPAFASIEGFAENLEATGVLGEGVTTADWTVP	257
QY	296	FMPAVIRSAFTWKGIL	310
DB	258	AWLDPT-----WOGI	267

RESULT 2
US-09-029-603-2
Sequence 2, Application US/09029603
Patent No. 6210935

GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Engel, Natalie
APPLICANT: Bietenhader, Jurg
APPLICANT: Toupet, Christine
APPLICANT: Pospishech, Andreas
TITLE OF INVENTION: Staurosporin Biosynthesis Gene Clusters
FILE REFERENCE: 4-20555/A/PCP
CURRENT APPLICATION NUMBER: US/09/029,603
CURRENT FILING DATE: 1998-03-20
EARLIER APPLICATION NUMBER: PCP/EP96/03643
EARLIER FILING DATE: 1996-08-19
NUMBER OF SEQ ID NOS: 11
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 2
LENGTH: 280
TYPE: PRT
ORGANISM: Streptomyces longisporoflavus
FEATURE:
OTHER INFORMATION: methyl transferase-like protein
US-09-029-603-2

Query Match 13.9%; Score 254; DB 4; Length 280;
Best Local Similarity 30.1%; Pred. No. 3.2e-19;
Matches 69; Conservative 44; Mismatches 84; Indels 32; Gaps 8;

OY 67 KGIAEFYDE--SSGLMENIMGDHMHGFDY--SDSTVSLSDHRAAIRMIOESLRFASV 121
DB 12 EAVGELYRLTLNANDSEFNENHIGYWDTPGSEATEEAMDRITDV-FIERLNAAYT- 69
OY 122 SEERKWKPSIYDVCGGIGSSRYLAKKFGATSVGTTTSPVOAQRANLAAAGLADRV 181
DB 70 -----SHVLDLCCGGGPELRYAVARTGAVTGISISEQIRFANRLAAEAGVADRAV 121
OY 182 FOVADALQPFSDGQFDIYVMSGEHMPKAKFYGLARVAPAAITIIYTWCHRDIGP 241
DB 122 FONGAMMLPFRADASFDAVMALESICHPDRQVTEVCYULRPGRIYLVLDIFERR--P 179
OY 242 DEOSLHP---WEDDLKIKICDVAIYLPAMCSTSDYVKLL--QSLSLADI 284
DB 180 RKAHMHGIDKFKRDLMTAD-----IDYVALLHRSGLRLREI 219

RESULT 3

US-09-041-718-5
Sequence 5, Application US/09041718A
Patent No. 6225075
GENERAL INFORMATION:
APPLICANT: Bard, Martin
TITLE OF INVENTION: DNA encoding sterol methyltransferase
FILE REFERENCE: 740.003051
CURRENT APPLICATION NUMBER: US/09/041,718A
CURRENT FILING DATE: 1998-03-13
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 363
TYPE: PRT
ORGANISM: Trilicium atrivum
US-09-041-718-5

Query Match 13.1%; Score 239; DB 4; Length 363;
Best Local Similarity 25.1%; Pred. No. 2e-17;
Matches 85; Conservative 49; Mismatches 99; Indels 106; Gaps 14;

OY 69 IAEFYDESSGLM-AIMGDHMH-HGFYDSSTVSLSDHRAAIRMIOESLRFASV 123
DB 70 VNKYYDLATSFYEGWGSFPAHRMNGESLRKHEHLALQLEL----- 117
OY 124 ERKMPKSIYDVCGGIGSSRYLAKKFGATSV-GITLSPVOAQRANLAAAGLADRV 182
DB 118 ---KPKMKVLDVCGGIGGPLERFA-RFSSTSVTGNNNDYQITRGKALNRSVIGATCDF 173

OY 183 OVADALQPFSDGQFDIYVMSGEHMPKAKFYGLARVAPAAITIIYTWCHRDIGP 235
DB 174 VNADEPMKPFSDGQFDIYVMSGEHMPKAKFYGLARVAPAAITIIYTWCHRDIGP 233
OY 236 ---HR-----DLG---PDEOSLHP-----WEDDLKIKICDVAIYLPAMC- 268
DB 234 NNATKRRIKDELELNGPLDISTRQCLQAVDAGEFVIMDQDLADSDPLPILPDSR 293
OY 269 -----TSDYVKLLQSLSDIDIKSEDMSPFVAPFPAVIRSAFTWKLSSL 313
DB 294 FSLSEFRLTYGRITIRNANVKLEYVGLAPESQSRVSSPL----- 333
OY 314 LSSGOKTKGALAPLMIEGYKKDL---IKFAITTCRK 349
DB 334 ---EKAAG-----LVEGCKKEIFTPMYPFVV--RKP 360

RESULT 4

US-09-036-987A-7
Sequence 7, Application US/09036987A
Patent No. 6143526

GENERAL INFORMATION:
APPLICANT: Balzer, Richard H.
APPLICANT: Broughton, Mary C.
APPLICANT: Crawford, Kathryn P.
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Merlo, Donald J.
APPLICANT: Treadway, Patil J.
APPLICANT: Turner, Jan R.
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dow Agrosciences LLC Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,987A
FILING DATE: 09-MAR-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stuart, Donald R.
REGISTRATION NUMBER: 28,479
REFERENCE/DOCKET NUMBER: 50,608
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317)337-4816
TELEFAX: (317)337-4847

INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 275 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-036-987A-7

Query Match 12.0%; Score 220.5; DB 4; Length 275;
Best Local Similarity 25.9%; Pred. No. 1.3e-15;
Matches 76; Conservative 47; Mismatches 126; Indels 45; Gaps 9;

OY 69 IAEFYDESSGLMENIMGD--HMHGFYDSSTVSLSDHRAAIRMIOESLRFASVSEERS 126
DB 12 VGVYDVLVPLNLSVAGPCAIHHGYWENDGRASWQ-----QAADLTLTDLVART 61
OY 127 KMPKSI--VDVCGGIGSSRYLAKKFGATSVGTTTSPVOAQRANLAAAGLADRV 184


```

Db 141 TGCNTVIGLNNNDYQAKATYKAYKYLSDQMDVKGDFPKMDFEETFDKVALEXTCHA 200
Oy 210 PKKAFVGLARVAAPGALIIIVTWCNRDLGPDQSLHFWEDLLKICDARY-----L 263
Db 201 PKLEGYSEIYKYLKRGCFAYEWMTDKDENNEH-----RKI--AVEIELGDGI 251
Oy 264 PWCSTSDVKLLQSLSDIKSED-----MSRFAPFWPAY-----IRSAF 305
Db 252 PKMFHVDAVKALKKNGCFVLVSEDLADNDEIPWYPTLTGEMKVVQNLANLATEFFRTSY 311
Oy 306 TWKGLSS-----LLSGOKTIGAL---AMPLMIEGYKKDLIKFAITTCRKE 350
Db 312 LGROFTTANVTYMEKLGILAPBESKEYTALLENAAVGLVAGSKSLFTPMPLVARKPE 369

```

RESULT 8 US-09-041-718-2

```

; Sequence 2, Application US/09041718A
; Patent No. 6225075
; GENERAL INFORMATION:
; APPLICANT: Bard, Martin
; TITLE OF INVENTION: DNA encoding sterol methyltransferase
; FILE REFERENCE: 740.003US1
; CURRENT APPLICATION NUMBER: US/09/041,718A
; CURRENT FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-041-718-2

```

```

Query Match 10.3%; Score 189; DB 4; Length 376;
Best Local Similarity 21.6%; Pred. No. 5,7e-12;
Matches 74; Conservative 55; Mismatches 133; Indels 80; Gaps 13;

```

```

Oy 59 KDDKKRLQKIAE-----FDDESSGLMENWGMHMG-FYDSDS---TVSLSDHRAAQIR 110
Db 59 KDDKKRLNDYSQLTGHHYNNLWTFEYGVGSSFFHSRYKKEAFQATARHGHFAHQM 118
Oy 111 MIOESLRFASVSEERSKWPFSIVDVCIGSSSRYLAKFGATSVGITLSPVOAQRANL 170
Db 119 NINENMK-----VLDVCGVGGPGRETRTTRTDEIYGLNNNDYQIERANHY 164
Oy 171 AAAGLADKVSFOVADALQPFSDQFDLWMSGESGEMPDRAKFGVGLARVAAPGALII 230
Db 165 AKKYHLDRKLSYKGDPMDFEPESFDVAIVAEATVHAPVLEGVYSEIYKYLKPGCIG 224
Oy 231 IYTWK-----HRLG-----PDEQSLH--PWEDDLKIKICA- 260
Db 225 VEWMTDKYDETNHEHRIAYGIEVGDSIPKWSRKVAEOALKNGFEIEYOKDLADVD 284
Oy 261 ----YILP-----AMCST-SDYVKLLQSLSDIKSEDSMSRVAPFWPAYISAFWKGL 310
Db 285 DEIPWYPTLSDGLKFCQFTGDIYVPRISRI-----GRYITTESVGLMEKI----- 330
Oy 311 SSLSSSGOKTIGAL---AMPLMIEGYKKDLIKFAITTCRKP 349
Db 331 -GLAPKSGKQVTHALEDAVNVLVEGGRQKLTFTPMPLVYVARKP 371

```

RESULT 9 US-09-036-987A-13

```

; Sequence 13, Application US/09036987A
; Patent No. 6143526
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Merlo, Donald J.
; APPLICANT: Treadway, Patli J.

```

```

; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; TITLE OF INVENTION: Production
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dow Agrosciences LLC Patent Department
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,987A
; FILING DATE: 09-MAR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stuart, Donald R.
; REGISTRATION NUMBER: 28,479
; REFERENCE/DOCKET NUMBER: 50,608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317)337-4816
; TELEFAX: (317)337-4847
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 283 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-036-987A-13

```

```

Query Match 8.3%; Score 152.5; DB 4; Length 283;
Best Local Similarity 24.2%; Pred. No. 3,3e-08;
Matches 48; Conservative 28; Mismatches 77; Indels 45; Gaps 4;

```

```

Oy 85 GDHMHGFYD-----SDSTVSLSDHRAAQIRMIQESLRFASVSEERSKWPFSI 132
Db 11 GPRLHGHWAGCYREDACATPMSDADQLTD-----LFLDKALIRGA-----HL 55
Oy 133 VDVCGIGSSSRYLAKFGATSVGITLSPVOAQRANLAAAGLADKVSFOVADALQPF 192
Db 56 FDLGCGNQPVYRAACASCVRTGITVNAQHLLAAATRLANETGLAGSLFEDLVDAQLPY 115
Oy 193 SDGQFDLWMSGESGEMPDRAKFGVGLARVAAPGALIIIVTWCNRDLGPDQSLHFWED 252
Db 116 PDGFPQAMAMQSVQVQIDQAAAIREVHRILEPGRFVLDIITRVRLPEE----- 166
Oy 253 LTKKICDAVYLPWCSTS 270
Db 167 -----YAAVWTGTT 175

```

RESULT 10 US-09-370-700-13

```

; Sequence 13, Application US/09370700
; Patent No. 6274350
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patli J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50489 DIV1
; CURRENT APPLICATION NUMBER: US/09/370,700
; CURRENT FILING DATE: 1999-08-09

```

```

      ORGANISM: Escherichia coli
US-08-457-245-9

Query Match          8.3%; Score 152; DB 1; Length 282;
Best Local Similarity 24.1%; Pred. No. 3,7e-08;
Matches 61; Conservative 40; Mismatches 94; Indels 58; Gaps 12;

QY   71 EFYDESSSLMENVNDHHNG--FYDSDSVYSLSDBRAADIRMIOESIRFASVSEBSK 127
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    18 EHYDLGNDLFGRMDLPDFEQYSCAYWKAD--NLESAQQAKLKMKCEKLQL-----K 66

QY   128 WPKSIVDGGCGIGSSRRYLAKKFGATSGITLSPVOAORANALAAAGSLADKVSFOVADA 187
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    67 PGMRYLDIGCCWGGGLAHNAHSNYDVSYGVITISAQOKMAQ--ERCSEGIDVTI----- 117

QY   188 LQGPFS--GGFDLVSMNSEGEHMPDK--ANFVGELARVAAPGAIIITYTWCNRDLGPDE 243
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    118 LLQDYRDLDNQEDFRIVSYGMFEHWGPKKNVDYFAVVDRNLKEGIFLLHT-----IGSKK 172

QY   244 QSLA--PWEODLLKICCAYYILPAMCSTSDYVKLLQSLSLODIKSEDMSRFVAF----- 296
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    173 TDLNVDPW-----INKYIFPNCLPS--VRQIAQSEPHFTVMEDWHNFAGADYDTTLM 222

QY   297 -----WPAP 300
      || :
Db    223 AWYERFLAMPET 235

RESULT 12
US-09-355-166-20
Sequence 20, Application US/09355166
Patent No. 6316241
GENERAL INFORMATION:
APPLICANT: Genencor International, Inc.
TITLE OF INVENTION: Alpha/Beta Hydrolase-Fold Enzymes
FILE REFERENCE: GC511-PCT
CURRENT APPLICATION NUMBER: US/09/355,166
CURRENT FILING DATE: 1999-07-20
NUMBER OF SEQ ID NOS: 21
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 20
LENGTH: 233
TYPE: PRP
ORGANISM: Bacillus
US-09-355-166-20

Query Match          5.9%; Score 107.5; DB 4; Length 233;
Best Local Similarity 28.1%; Pred. No. 0.0018;
Matches 45; Conservative 24; Mismatches 74; Indels 17; Gaps 6

QY   133 IVDVCGIGSGSSRYLAKKFGATSGITLSPVOAORANALAAAGLADKVSFOVADALOOP 191
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    39 IIDACGCIGTGTAAYLGHLHPYTV-VDRKPIMLEKAKKRFANEGLA--IPAYQAELENRP 95

QY   192 FSDGQFDLVSMNSEGEHMPDKAKFVGEELARVAAP-GAIIITYTWCNRDLGPDEQSILRME 250
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    96 FSSESFSCVLS-ESVLSFSRLTSLOEISRVLYKPCMGMLGIEALKKKPPPAEK----- 148

QY   251 QDLKKICDAYYLPAMCSTSDYVKLLQSLSLODIKSEDM 290
      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    149 ----KOMDFGYPTCLHESEWHKLRISTGFQ--KTEAMS 182

RESULT 13
US-08-576-626A-56
Sequence 56, Application US/08576626A
Patent No. 5998194
GENERAL INFORMATION:
APPLICANT: Summers, R.G.
APPLICANT: Katz, L.
APPLICANT: Donadio, S.
APPLICANT: Staver, M.J.
TITLE OF INVENTION: POLYKETIDE-ASSOCIATED SUGAR

```

```

1  TITLE OF INVENTION: BIOSYNTHESIS GENES
2  NUMBER OF SEQUENCES: 60
3  CORRESPONDENCE ADDRESS:
4  ADDRESSEE: Abbott Laboratories
5  STREET: 100 Abbott Park Road
6  CITY: Abbott Park
7  STATE: Illinois
8  COUNTRY: USA
9  ZIP: 60064-3500
10 COMPUTER READABLE FORM:
11 MEDIUM TYPE: Diskette
12 COMPUTER: IBM Compatible
13 OPERATING SYSTEM: DOS
14 SOFTWARE: FASTSEQ Version 2.0
15 CURRENT APPLICATION DATA:
16 APPLICATION NUMBER: US/08/576,626A
17 FILING DATE: 21-DEC-1995
18 CLASSIFICATION: 435
19 PRIOR APPLICATION NUMBER:
20 APPLICATION NUMBER:
21 FILING DATE:
22 ATTORNEY/AGENT INFORMATION:
23 NAME: Dianne Casuto
24 REGISTRATION NUMBER: P-40,943
25 REFERENCE/DOCKET NUMBER: 5857 US.O1
26 TELECOMMUNICATION INFORMATION:
27 TELEPHONE: (847) 938-3137
28 TELEFAX: (847) 938-2623
29 TELEX:
30 INFORMATION FOR SEQ ID NO: 56:
31 SEQUENCE CHARACTERISTICS:
32 LENGTH: 237 amino acids
33 TYPE: amino acid
34 STRANDEDNESS: single
35 TOPOLOGY: linear
36 MOLECULE TYPE: NO. 5998194e
37 OS-08-576-626A-56

```

```

Query Match          5.7%;  score 104.5;  DB 2;  length 237;
Best Local Similarity 27.0%;  Pred. No. 0.004;
Matches 48;  Conservative 20;  Mismatches 57;  Indels 53;  Gaps 11;

OY      68  GIAEYDESSGLMENIWGDHMHGFFYDSDSTVSLSDH--RAAQI-RMIOESLRFASVSEE 124
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      5  GFAEYIDR-----FYRRSG--KDYAAEAQVARIYDRDRPSPAS--- 40

OY      125 RSKMKSIVDYCGGIGGSSRYLAKKEGATSVGITSLPVOGAOKA-----NALAAQGLADKV 180
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      41  -----SLIDVACGTGTHLRFPADLFDDYT-GLELSAAMIEVARPOLGIPVLQG--DWR 91

OY      181 SFOVADALQOPFSDQFDPLVWSMESG-EHMPDKAF--VGLERAAAPGAILITVT 234
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      92  DFAL-----DREFAVTCTMFSISGHMRGAEIDQALSFARHLAPGAVVVEPW 140

RESULT 14
US-08-576-626A-32
; Sequence 32, Application US/08576626A
; Patent No. 5998194
; GENERAL INFORMATION:
; APPLICANT: Summers, R.G.
; APPLICANT: Katz, L.
; APPLICANT: Donadio, S.
; APPLICANT: Scaver, M.J.
; TITLE OF INVENTION: POLYKETIDE-ASSOCIATED SUGAR
; TITLE OF INVENTION: BIOSYNTHESIS GENES
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: Illinois
; COUNTRY: USA
;

```

```

ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/576,626A
FILING DATE: 21-DEC-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dianne Casuto
REGISTRATION NUMBER: P-40,943
REFERENCE/DOCKET NUMBER: 5857, US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (847) 938-3137
TELEFAX: (847) 938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 2544 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: NO. 5998194e
US-08-576-626A-32

```

```

Query Match          5.7%; Score 104.5; DB 2; Length 2544;
Best Local Similarity 27.0%; Pred. No. 0.16;
Matches 48; Conservative 20; Mismatches 57; Indels 53; Gaps 11.

QY 68 GAAFYDESSGLMENIMGDHMHGHYDSDSYLSLDH--RAQOI-RMIOESLRFASVSE 124
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 742 GAELLYD-----FYRGRG--KDYAAEAQAVRLVDRPLPSAS---- 777
QY 125 RSKPKPSIVDYGCGIGGSSRYLAKKFGATSVGITLSPVOAQRA---NALAAOGIADKY 180
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 778 -----SLIDVACGIGTHLRFADLFDVLT-GLELSAAMIEVARPOLGIPVLQG--DMR 828
QY 181 SFQYADALQCFPSDGOEFLVWSMMSG-EMHPDKAF---YGLARVAAPGALIIIVTV 234
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 829 DFL-----DREPDATVCMFSSIGHMRGAEILDALASFARHLAAGVVVEPW 877

RESULT 15
US-08-401-068-12
; Sequence 12, Application US/08401068
; Patent No. 5859335
;
GENERAL INFORMATION:
APPLICANT: Patton, David
TITLE OF INVENTION: Enhanced Biotin Biosynthesis in Plant Tissue
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/401,068
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/351,970

```


GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 17:38:15 ; Search time 15 Seconds

(without alignments)
1075.482 Million cell updates/sec

Title: US-09-857-613A-28

Perfect score: 1830
Sequence: 1 MATVVRIPITISCHIRIFRS.....IEGKKDLKFAITCKRPE 350

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PC1_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/PC1US_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	225	12.3	344	10 US-09-779-144A-7	Sequence 7, Appli
2	212.5	11.6	361	10 US-09-779-144A-4	Sequence 4, Appli
3	203.5	11.1	383	10 US-09-801-368-96	Sequence 96, Appli
4	203.5	11.1	383	10 US-09-779-144A-2	Sequence 2, Appli
5	141	7.7	251	9 US-10-260-877-54	Sequence 54, Appli
6	139	7.6	451	9 US-09-738-626-4136	Sequence 4136, Ap
7	139	7.6	451	10 US-09-845-248-2	Sequence 2, Appli
8	117.5	6.4	206	9 US-09-738-626-4923	Sequence 4923, Ap
9	112	6.1	251	9 US-09-738-626-6365	Sequence 6365, Ap
10	110	6.0	218	9 US-09-738-626-4999	Sequence 4999, Ap
11	107.5	5.9	233	10 US-09-950-368-10	Sequence 20, Appli
12	105	5.7	256	10 US-09-815-242-12004	Sequence 12004, A
13	105	5.7	348	10 US-09-939-521-11	Sequence 11, Appli
14	99	5.4	241	10 US-09-815-242-12602	Sequence 12602, A
15	98.5	5.4	234	10 US-09-815-242-5314	Sequence 5314, Ap
16	97	5.3	266	12 US-10-007-693-66	Sequence 66, Appli
17	95.5	5.2	255	9 US-09-866-050A-706	Sequence 706, Appli
18	94.5	5.2	273	10 US-09-816-714-4	Sequence 4, Appli
19	94.5	5.2	680	9 US-09-738-626-4877	Sequence 4877, Ap

20	92	5.0	7968	9 US-10-077-130-5	Sequence 5, Appli
21	91.5	5.0	2201	9 US-10-085-476-2	Sequence 2, Appli
22	91.5	5.0	2201	12 US-10-029-907-3	Sequence 3, Appli
23	90.5	4.9	828	10 US-09-816-685-4	Sequence 4, Appli
24	90	4.9	1832	9 US-10-014-717-4	Sequence 4, Appli
25	89.5	4.9	257	9 US-09-738-626-6656	Sequence 6656, Ap
26	89	4.9	2894	10 US-09-941-611-23	Sequence 23, Appli
27	89	4.9	3011	10 US-09-916-359-2	Sequence 2, Appli
28	88.5	4.8	448	10 US-09-939-521-7	Sequence 7, Appli
29	88	4.8	207	10 US-09-921-397-94	Sequence 94, Appli
30	88	4.8	3011	9 US-09-747-419-20	Sequence 20, Appli
31	88	4.8	3011	9 US-09-891-894-3	Sequence 3, Appli
32	88	4.8	3011	9 US-09-995-937-20	Sequence 20, Appli
33	88	4.8	3011	10 US-09-742-659-4	Sequence 4, Appli
34	88	4.8	3011	10 US-09-738-626-4027	Sequence 9, Appli
35	88	4.8	3011	10 US-09-952-572-9	Sequence 2, Appli
36	88	4.8	3012	9 US-09-995-937-2	Sequence 2, Appli
37	88	4.8	3012	10 US-09-238-076-2	Sequence 164, App
38	87.5	4.8	281	9 US-09-984-245-164	Sequence 4, Appli
39	87	4.8	3782	9 US-09-860-846-4	Sequence 4, Appli
40	87	4.8	3782	10 US-09-861-289-4	Sequence 93, Appli
41	86.5	4.7	208	9 US-09-921-397-93	Sequence 4027, Ap
42	86.5	4.7	230	9 US-09-738-626-4027	Sequence 6979, Ap
43	86.5	4.7	692	9 US-09-738-626-6979	Sequence 16, Appli
44	85.5	4.7	276	12 US-10-033-078-16	Sequence 11636, A
45	85.5	4.7	315	10 US-09-815-242-11636	

ALIGNMENTS

RESULT 1									
US-09-779-144A-7									
Sequence 7, Application US/09779144A									
Patent No. US20020148006A1									
GENERAL INFORMATION:									
APPLICANT: Nes, David W.									
TITLE OF INVENTION: Transgenic Plants with Modified Sterol Compositions									
FILE REFERENCE: 11899, 0198, DVUS01 MOB:198									
CURRENT APPLICATION NUMBER: US/09/779,144A									
CURRENT FILING DATE: 2001-02-08									
PRIOR APPLICATION NUMBER: US 09/106,926									
PRIOR FILING DATE: 1998-06-29									
NUMBER OF SEQ ID NOS: 16									
SOFTWARE: PatentIn version 3.0									
SEQ ID NO 7									
LENGTH: 344									
TYPE: PRT									
ORGANISM: Zea mays									
US-09-779-144A-7									
Query Match									
Best Local Similarity 25.4%; Pred. No. 3.6e-14;									
Matches 77; Conservative 40; Mismatches 112; Indels 74; Gaps 9;									
12.3%; Score 225; DB 10; Length 344;									
QY	69	IAEFTDESSGLMENITWGDHMHGFFDOSDTSVLSLDRRAQIMQESLRFASVSEBRKW	128						
DB	51	VAKYVDLATSFEYWGSEFHFRAHMGESLRESIKRHHFLALQGLT-----KP	100						
QY	129	PSIYDVGGICGSSRYLTKKFGATSV-GITLSPVQAQANLAAAGLADKVSQVADA	187						
DB	101	GKRVLDVGGIGGIGPLRELA-RFSSVSVTLNNNEYQITGKELNRLAGISGCDPFAKDF	159						
QY	188	LOQPSDSCGFDLVWMSGESGHPDRAKEFVGEIARVAAPGAILIIVTW-----H	236						
DB	160	KMPPFDNNTFADVAVVLEACHAPDPVGCCKEILRYLPPQCAVYEMCTDHYDPNNATH	219						
QY	237	R-----DLG-----PDQSLHP-----MEQILKIKICDAVYLPAMCS-----	268						
DB	220	KRIKDEIEIGNGLPIRSTRQCLRAVKDAGFEVVDKDLAEDSPLPWYLPIDPSRFLSS	279						
QY	269	-----TSDYKILQSLSLQDLSKSDMSRFAPAFPRPAVIRSAFTWKGLSLLSSQ	318						

Db 280 FRUTSVGRMTPTWVKALEYGLADPOGSESSFILE-----KAAEGIVEGCK 326
 Oy 319 KTI 321
 Db 327 KEI 329

RESULT 2

US-09-779-144A-4
 ; Sequence 4, Application US/09779144A
 ; Patent No. US20020148006A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Nes, David W.
 ; TITLE OF INVENTION: Transgenic Plants with Modified Sterol Compositions
 ; FILE REFERENCE: 11899.0198. DIVS01 MOB7.198
 ; CURRENT APPLICATION NUMBER: US/09/779,144A
 ; PRIOR FILING DATE: 2001-02-08
 ; PRIOR APPLICATION NUMBER: US 09/106,926
 ; PRIOR FILING DATE: 1998-06-29
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 361
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 US-09-779-144A-4

Query Match 11.6%; Score 212.5; DB 10; Length 361;
 Best Local Similarity 25.4%; Pred. No. 6.6e-13;
 Matches 80; Conservative 54; Mismatches 144; Indels 37; Gaps 11;

Oy 55 QKPKK-DKKKKLQKGLAEFYDESSGLMENIWDHMHGFDSDSTVLSLSD-----HRAOI 109
 Db 59 RRKELETAKEVDFVDFETNLVTDIYEMGSGSFH--FSPSTPGSKHRTATRLHEMAY 116
 Oy 110 RMIOESLRFAVSSEERSKMKPSIVDVGCGIGSSRYLAKFGATSVGITLSPVOAORANA 169
 Db 117 DLIO-----VKPQKILIDVCGGVGPMRAIASRANVGVITINEYOVNARL 164
 Oy 170 LAAAGLADKVSFOVADALQPPSDGOFDLYWMESEHNPDKAKFYGELARVAAPRAIT 229
 Db 165 HNKAGIADALCEYVCGFLOMPDMSFDGAYSIKATCAHPKLEEVYAEIYRLKPGSMY 224
 Oy 230 IYTWCHRDGPRDEOSLHPMEODLKKICDAAYILPAMCSTSDYKILQSLQSLQDIKESDW 289
 Db 225 VSTEWTTEKFAEDDEH---VEYIOGIERGDALPGLRAYVDIAETKAKYGFELVRKDL 281
 Oy 290 SRPVA-PFMPAVI--RSAPFWKG--LSLSLSSGOKTIKALAMPIMI-----EGYKK 336
 Db 282 ASPPAEPMWTRLMKGRILAY-WRNHIVQIILSAGVAPKGVIVHEMLFKTADCLTRGET 340
 Oy 337 DLIK-PAITTCRKE 350
 Db 341 GIFSPMHMILCRKE 355

RESULT 3

US-09-801-368-96
 ; Sequence 96, Application US/09801368
 ; Patent No. US20020128250A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Busby, Robert
 ; APPLICANT: Hecht, Peter
 ; APPLICANT: Holtzman, Doug
 ; APPLICANT: Madden, Kevin
 ; APPLICANT: Maxon, Mary
 ; APPLICANT: Milne, Todd
 ; APPLICANT: No. US20020128250A1man, Thea
 ; APPLICANT: Royer, John
 ; APPLICANT: Salama, Sofie
 ; APPLICANT: Sherman, Amir
 ; APPLICANT: Silva, Jeff

;; APPLICANT: Summers, Eric
 ;; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fu
 ;; FILE REFERENCE: 109272.147
 ;; CURRENT APPLICATION NUMBER: US/09/801,368
 ;; CURRENT FILING DATE: 2001-03-07
 ;; PRIOR APPLICATION NUMBER: US 09/487,558
 ;; PRIOR FILING DATE: 2000-01-19
 ;; PRIOR APPLICATION NUMBER: US 60/160,587
 ;; PRIOR FILING DATE: 1999-10-20
 ;; NUMBER OF SEQ ID NOS: 440
 ;; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 96
 ; LENGTH: 383
 ; TYPE: PRT
 ; ORGANISM: Saccharomyces cerevisiae
 US-09-801-368-96

Query Match 11.1%; Score 203.5; DB 10; Length 383;
 Best Local Similarity 21.8%; Pred. No. 5.5e-12;
 Matches 78; Conservative 58; Mismatches 149; Indels 73; Gaps 12;

Oy 42 SAASERGEIVLEQKPKKDKKKLQKGLAEFYDESSGLMENI-----WGDHMHG-F 92
 Db 36 NSAKKEAVOKYLRNMDGRTDKDAEERL-EDYNETHSYNYVTDTYEGWGSSPFSRF 94
 Oy 93 YSDS---TVSLDHRRAOIRMOESLRFAVSSEERSKMKPSIVDVGCGIGSSRYLAK 149
 Db 95 YKGSFASASIRHEHLYAVKAGIORG-----DLVDVCGGVGAPAREIARF 140
 Oy 150 FGATSVGITLSPVOAORANLAAAGLADKVSFOVADALQPPSDGOFDLYWMESEHNP 209
 Db 141 TGCNVIIGLNNDQIQAAYAKKYNLSDDMDPKGDMDEEMTFDVVAIEATCA 200
 Oy 210 PDKAKFVGLARVAPGATIIIVTWCHRDGPRDEOSLHPMEODLKKICDAAY-----L 263
 Db 201 KLEGVSEIYKYLAKGCGFVAYEWMTKYDENNEH-----KRI-AYEELQDGI 251
 Oy 264 PAMCSTSDYKILQSLQSLQDIKESD-----WSRFVAPMPAV-----IRSAF 305
 Db 252 PKMFHVDVAKALKNGCFEVLVSEDLADNDEIFPWYPLTGEMKYVONLANLATFFFTSY 311
 Oy 306 TWKGLSS-----LLSSGOKTIKAL--AMPLMIGYKKDLKFAITTCRKE 350
 Db 312 LGRHFTTAMTVMEKILGABEGSKEVYALLENAAVGLAGSKSLFTPMILVAPRKE 369

RESULT 4

US-09-779-144A-2
 ; Sequence 2, Application US/09779144A
 ; Patent No. US20020148006A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Nes, David W.
 ; TITLE OF INVENTION: Transgenic Plants with Modified Sterol Compositions
 ; FILE REFERENCE: 11899.0198. DIVS01 MOB7.198
 ; CURRENT APPLICATION NUMBER: US/09/779,144A
 ; PRIOR FILING DATE: 2001-02-08
 ; PRIOR APPLICATION NUMBER: US 09/106,926
 ; PRIOR FILING DATE: 1998-06-29
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 383
 ; TYPE: PRT
 ; ORGANISM: Saccharomyces cerevisiae
 US-09-779-144A-2

Query Match 11.1%; Score 203.5; DB 10; Length 383;
 Best Local Similarity 21.8%; Pred. No. 5.5e-12;
 Matches 78; Conservative 58; Mismatches 149; Indels 73; Gaps 12;

Oy 42 SAASERGEIVLEQKPKKDKKKLQKGLAEFYDESSGLMENI-----WGDHMHG-F 92
 Db 36 NSAKKEAVOKYLRNMDGRTDKDAEERL-EDYNETHSYNYVTDTYEGWGSSPFSRF 94


```

QY 93 YDSOS---TVSLGHPRAQIMEDISLRPASVSEBSKMKPSIYDVCCGIGGSSRYLAK 149
    | : : : : |
Db 95 YKGSFPAASIAIRNHXYIAIKRIGDRC-----DLVIDVCCGIGGFAREIARF 140
    | : : : : |

QY 150 FGATSVGTTLSPVQAQRANMLAAAGLADRVSEFQVADALQOPSDGQFDLIVMSGESGEM 209
    | : : : : |
Db 141 TGCVAVICGLNNNDXYIAAKAYAKRYNLSDDOMDFKCGPMKMDFEENFPDRVYAIEATCHA 200
    | : : : : |

QY 210 PDKAFVQELARVAAPGIIIIIVWCHRDLDGPDQSLHPMKDILKTKICQAY-----L 263
    | : : : : |
Db 201 PKLEGVYSEIKRYLKPQGTFRVYEWMTDKYDENNPH-----KRI--AYELDLDGI 251
    | : : : : |

QY 264 PAMCSTDYVLLQSLSLQDIKSD-----WSRFVAPFWPVA-----IRSAF 305
    | : : : : |
Db 252 PKMFHVDAARFALKNCGFVLYVSDLDLNDDEIPIWYPTLIGEMKYVONLANLATFFRTSY 311
    | : : : : |

QY 306 TWKGLSS-----LLSGOKTIGAL---AMPIMIGYKKDLKFAIITOCRKE 350
    | : : : : |
Db 312 LGRQFTTAMVYMEKLGIAPEGSSEVYRANALENAAVGLVAGCKSKLTFPMMLFPAARKE 369
    | : : : : |

```

```

RESULT 5
US-10-260-877-54
; Sequence 54, Application US/10260877
; Publication No. US20030021813A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Chovan, Linda E.
; APPLICANT: Hessler, Paul E.
; APPLICANT: Reich, Karl A.
; TITLE OF INVENTION: ESSENTIAL BACTERIA GENES AND GENOME
; TITLE OF INVENTION: SCANNING IN HAEMOPHILUS INFLUENZAE FOR THE IDENTIFICATION OF
; FILE REFERENCE: 6565_US.P1
; CURRENT APPLICATION NUMBER: US/10/260,877
; PRIOR FILING DATE: 2002-09-30
; PRIOR FILING DATE: 2000-08-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 54
; LENGTH: 251
; TYPE: PRT
; ORGANISM: H. Influenzae
US-10-260-877-54

Query Match          7.7%: Score 141; DB 9; Length 251;
Best Local Similarity 25.0%; Pred. No. 4.4e-06;
Matches 47; Conservative 37; Mismatches 74; Indels 30; Gaps 5;

QY   130 KSIVVGGGIGSSRYLAKKFGATSVGITLSPVQAQRANALAAOGIADKVSFOVDALO 189
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    40 KKVLEAACMGTTAIGLKAFQFCHEGVLDENALAKAANIEANKLOEKIHQRANAK 99
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY   190 QPFSGDGFDLVSMSESGHP--DKAFVGLARVAAPCAIIITYWCHRDLGPDDQSILH 247
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    100 LPFEDESDGIIVNEAMLFWLPVEARKKAIAEYFVLKPNGLL--THDVMLVGNDHQIT- 156
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY   248 PWEDLLKKICDVIYILPMACSTSDYVKLLQSLDIDIKSDDMSRFAPMPAVIRSAFTW 307
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    157 --LENNRAAI-----NVTVPPLTKDGWKGI---FOESGEFRNVDTF 191
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY   308 KGLSSILS 315
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    192 SCEMTLLS 199

RESULT 6
US-09-738-626-4136
; Sequence 4136, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
```

```

; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SEMOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4136
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-4136

```

```

Query Match          7.6%: Score 139; DB 9; Length 451;
Best Local Similarity 22.3%: Pred. No. 1.6e-05;
Matches 63; Conservative 42; Mismatches 122; Indels 56; Gaps 12.

QY 53 LEQRP-----KKDKKIKLQKGAIEFYDESSGLMENIWDGHH--GEYDSSTV 99
DB 143 MEQAPAMRRKALLINGLASHSRKSRDKALSYHNDGNEYSILFLDSMTYTCAYTPPEPS- 201
QY 100 SLSDHRAAOIMIOESLRFPAVSSEERSKWPKSIVYVGGIGSSRYLAKKFGATSVGITL 159
DB 202 SLEEQEAKRYRLIEFKRL-----KEGDRLLDVGCGMGGMVTR-AAKHGVKALGVTL 252
QY 160 SPVQANRANATLAAAGLADKVSFYVADALQOPFSDGOEFLVMSMESGEHM-----PDKAK 214
DB 253 SEQQEWGQAEIKRGLEDLAEIRMDYRDVP--ETGPDALSAIGLIEHIGVNNYPDYFE 310
QY 215 FVGLARVAAPACATIIITYTWCNRDLGCPPEQSHHPEDOLIKTCATYLP--AKMSTSDY 272
DB 311 LLSSKLKT---GGMLL---NHSITYPNRPRH-----AGAFIDRYIFPDDELGSGTL 357
QY 273 VKLLQSLSLQDIKSED-----MSRFAPFMPAVIRSA 304
DB 358 IKHMQDNGFEVILHEENLRFQYRTIHANCENLKEMMEAEVELA 400

RESULT 7
US-09-845-248-2
: Sequence 2, Application US/09845248
: Patent No. US20020042106A1
: GENERAL INFORMATION:
: APPLICANT: Nampoothiri, Madhavan
: TITLE OF INVENTION: Nucleotide Sequences Which Code for the CMA Gene
: FILE REFERENCE: 032301 WD 1161
: CURRENT APPLICATION NUMBER: US/09/845,248
: CURRENT FILING DATE: 2001-08-30
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 2
: LENGTH: 451
: TYPE: PRT
: ORGANISM: Corynebacterium glutamicum
US-09-845-248-2

Query Match          7.6%: Score 139; DB 10; Length 451;
Best Local Similarity 22.3%: Pred. No. 1.6e-05;

```



```

: TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
: FILE REFERENCE: 249-125
: CURRENT APPLICATION NUMBER: US/09/738, 626
: CURRENT FILING DATE: 2000-12-18
: PRIOR APPLICATION NUMBER: JP 99/377484
: PRIOR FILING DATE: 1999-12-16
: PRIOR APPLICATION NUMBER: JP 00/159162
: PRIOR FILING DATE: 2000-04-07
: PRIOR APPLICATION NUMBER: JP 00/280988
: PRIOR FILING DATE: 2000-08-03
: NUMBER OF SEQ ID NOS: 7059
: SOFTWARE: PatentIn ver. 3.0
: SEQ ID NO 4999
: LENGTH: 218
: TYPE: PRT
: ORGANISM: Corynebacterium glutamicum
: US-09-738-626-4999

```

[illegible]

```

RESULT 11
US-09-950-368-20
: Sequence 20, Application US/09950368
: Patent No. US20020061580A1
: GENERAL INFORMATION:
: APPLICANT: Genencor International, Inc:
: TITLE OF INVENTION: Alpha/Beta Hydrolase-Fold Enzymes
: FILE REFERENCE: GC511-PC7
: CURRENT APPLICATION NUMBER: US/09/950,368
: CURRENT FILING DATE: 2001-09-10
: PRIOR APPLICATION NUMBER: 09/355,166
: PRIOR FILING DATE: 1999-07-20
: NUMBER OF SEQ ID NOS: 21
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 20
: LENGTH: 233
: TYPE: PRT
: ORGANISM: Bacillus
: US-09-950-368-20

```

[illegible]

```

RESULT 12
US-09-815-242-12084
; Sequence 12084, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:

```

```

APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIORITY APPLICATION NUMBER: 60/191,078
PRIORITY FILING DATE: 2000-03-21
PRIORITY APPLICATION NUMBER: 60/206,848
PRIORITY FILING DATE: 2000-05-23
PRIORITY APPLICATION NUMBER: 60/207,727
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: 60/242,578
PRIORITY FILING DATE: 2000-10-23
PRIORITY APPLICATION NUMBER: 60/253,625
PRIORITY FILING DATE: 2000-11-27
PRIORITY APPLICATION NUMBER: 60/257,931
PRIORITY FILING DATE: 2000-12-22
PRIORITY APPLICATION NUMBER: 60/269,308
PRIORITY FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12084
LENGTH: 256
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
IS-09-815-242-12084

```

[illegible]

```

: RESULT 13
: US-09-939-521-11
:
: Sequence 11, Application US/09939521
: Patent No. US20020119466A1
: GENERAL INFORMATION:
: APPLICANT: Meyers, Rachel
: APPLICANT: Williamson, Mark
: APPLICANT: Rudolph-Owen, Laura A.
: TITLE OF INVENTION: 46863, A NOVEL HUMAN METHYLTRANSFERASE
: TITLE OF INVENTION: AND USES THEREOF
: FILE REFERENCE: NMI-183
: CURRENT APPLICATION NUMBER: US/09/939,521
: CURRENT FILING DATE: 2001-08-24
: PRIORITY APPLICATION NUMBER: 60/227,867
: PRIORITY FILING DATE: 2000-08-24
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: FastSeq for Windows Version 4.0

```

SEQ ID NO 11
LENGTH: 348
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-939-521-11

Query Match 5.7% Score 105; DB 10; Length 348;

Best Local Similarity 23.6%; Pred. No. 0.025; Mismatches 74; Indels 58; Gaps 12;

Matches 51; Conservative 33; Mismatches 74; Indels 58; Gaps 12;

Db 103 DHRAAQIMIDSLRFS-----VSEERSKMPKSIIVDVGCGIGSSRYLAKKFGATSVGI 157

Db 28 DHYGHEMLQDTYRTLSYRNAIIONKDLFKDKIYLDVGCSTGLISMFAAKHGAHVIGV 87

Db 158 TLSPVOQRANALAAOGLAKVSE--QVADALQPFSDQDFDLYVSMESG-----EHM 209

Db 88 DMSI-TEMKELVELNFSKIDTLRGKLED-VHLPPP--KVDIISEMNGYFLLYESM 143

Db 210 PDKKFGVGLARVAAPGAIITIVTWCNRDLGPEOSLHPWEDDLKKICDAYY---LPA 265

Db 144 MDYVLYARD--HYLVEGGLIF-----PDKCSIH-----LAGLEDQYKDEKINY 185

Db 266 WCSTSDYVKLQSLQDISEDSMSEFVAPFMPAVI 301

Db 186 W-----QDVYGPDS-----PFVPLVL 202

RESULT 14

US-09-815-242-12602

Sequence 12602, Application US/09815242

Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes In

FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 12602

LENGTH: 241

TYPE: PRT

ORGANISM: Staphylococcus aureus

US-09-815-242-12602

Query Match 5.4% Score 99; DB 10; Length 241;

Best Local Similarity 18.6%; Pred. No. 0.058; Mismatches 108; Indels 100; Gaps 11;

Matches 58; Conservative 46; Mismatches 108; Indels 100; Gaps 11;

Db 54 EOKPKDDKKLQKIGAEFDESSGLMENIWGDHMHGFDSDSYVLSLDRRAQIOTMIO 113

Db 3 DNKANKQVHRVFPONISKYDRLNNI-----ISFEQHKVVRKRYMK 43

Db 114 E-SLRPASVSEERSKMPKSIIVDVGCGIGSSRYLAKKFGATS--VGITLSPVOQRANAL 170

Db 44 DMVGROGT-----KALDVCCGIGDWTIALSKANGPIGVYTGIDFSSNMILEYKREK 93

Db 171 AAAGLADKVSFOYADALQOPFSDGQFDLYVSMESGEHMPDKAFVGLARVAAPGAIIT 230

Db 94 TASH---ENVKLVGDAMLEPFEDNSFDYVITIGFGLRNPVLYALAKEMNRKLRPGMYV 150

Db 231 IVTWCNRDLGPEOSLHPWEDDLKKICDAYLYLAWCSTSYVVKLQSLQDISEDSM 290

Db 151 CLETSOPLT-----PVRKOMAYLYF----- 170

Db 291 REVAPFMP---AVIRSAFTMGSLSSLSGQTKR-----GALAMP 329

Db 171 KFMPIFGKLFRAKSEYEIMLQOSTFNPGRKELKRMEEAGFINVRVRSFTGGVAAHML 230

Db 330 MIEGK-KDLIK 340

Db 231 ---GYKENDNTR 239

RESULT 15

US-09-815-242-5314

Sequence 5314, Application US/09815242

Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes In

FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 5314

LENGTH: 234

TYPE: PRT

ORGANISM: Staphylococcus aureus

US-09-815-242-5314

Query Match 5.4% Score 98.5; DB 10; Length 234;

Best Local Similarity 20.6%; Pred. No. 0.062; Mismatches 75; Indels 35; Gaps 5;

Matches 37; Conservative 33; Mismatches 75; Indels 35; Gaps 5;

Db 54 EOKPKDDKKLQKIGAEFDESSGLMENIWGDHMHGFDSDSYVLSLDRRAQIOTMIO 113

Db 1 DNKANKQVHRVFPONISKYDRLNNI-----ISFEQHKVVRKRYMK 41

Db 114 E-SLRPASVSEERSKMPKSIIVDVGCGIGSSRYLAKKFGATS--VGITLSPVOQRANAL 170

Db 42 DMGVRGT-----KALDYCCGTGDMTIALSKAVGPIGEVTGIDFSENMLEYGKEK 91
Qy 171 AAAGLADYVSPQVADALQOPESDGGFDLWMSHESGEHMPDKAKFYGELARVAAPGAIIT 230
Db 92 TASM---ENVKLVHGDAMELPFEDNSFDYVTIGFGLRNVPDYLVALKEMNRVLKPGGMVY 148

Search completed: March 18, 2003, 17:42:40
Job time : 18 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 17:36:22 ; Search time 19 Seconds

(Without alignments)
1770.897 Million cell updates/sec

Title: US-09-857-613A-28

Perfect score: 1830

Sequence: 1 MATVRIPIPTICHIHFRS.....IEGKKDLIKFAITTCRKP 350

Scoring table: BLOSUM62

Gapop 10.0 , Gapect 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR_73:*

2: PIR1:*

3: PIR2:*

4: PIR3:*

5: PIR4:*

6: PIR5:*

7: PIR6:*

8: PIR7:*

9: PIR8:*

10: PIR9:*

11: PIR10:*

12: PIR11:*

13: PIR12:*

14: PIR13:*

15: PIR14:*

16: PIR15:*

17: PIR16:*

18: PIR17:*

19: PIR18:*

20: PIR19:*

21: PIR20:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1154.5	63.1	348 2	gamma-tocopherol m
2	659	36.0	2 AE2031	gamma-tocopherol m
3	647.5	35.4	317 2	hypothetical prote
4	392.5	21.4	330 2	hypothetical prote
5	392	21.4	AC2071	hypothetical prote
6	247	13.5	576226	probable sterol 24
7	241	13.2	T03845	probable sterol 24
8	238	13.0	T06795	probable sterol 24
9	227	12.4	T44579	C5-O-methyltransfe
10	226	12.3	T04138	sterol 24-C-methyl
11	225	12.3	UC6531	avermectin B 5-O-m
12	221.5	12.1	T01572	sterol 24-C-methyl
13	221.5	12.1	T03848	probable sterol 24
14	218.5	11.9	T39579	sterol methyltrans
15	212.5	11.6	T42375	probable sterol 24
16	210	11.5	S63686	sterol 24-C-methyl
17	204	11.1	T50969	probable DELTA(24)
18	203.5	11.1	T06780	probable sterol 24
19	194.5	10.6	S42003	sterol 24-C-methyl
20	190	10.4	T10173	sterol 24-C-methyl
21	188	10.3	T19153	protein T1N15.23 f
22	182	9.9	H96762	hypothetical prote
23	180	9.9	T33885	hypothetical prote
24	169	9.2	B70797	probable transfera
25	164.5	9.0	H84364	membrane protein f
26	163	8.9	C82951	conserved hypothet
27	161	8.8	T29330	hypothetical prote
28	161	8.8	C70549	hypothetical prote
29	157	8.6	A81324	cyclopropane-fatty
			G86977	probable methyltra

30	156.5	8.6	273	2	T34740	hypothetical prote
31	156	8.5	274	2	B70901	hypothetical prote
32	156	8.5	441	2	AF3437	cyclopropane-fatty
33	153.5	8.4	347	2	C70723	probable methyltra
34	153	8.4	382	2	AC0696	cyclopropane-fatty
35	152	8.3	382	2	AA4292	cyclopropane-fatty
36	152	8.3	382	2	AG0291	cyclopropane-fatty
37	150	8.2	382	2	B90925	cyclopropane fatty
38	150	8.2	382	2	F85773	cyclopropane fatty
39	149	8.1	306	2	S18533	eryg protein - Sac
40	149	8.1	409	2	D87426	cyclopropane-fatty
41	148.5	8.1	278	2	E82965	hypothetical prote
42	147.5	8.1	452	2	B75306	probable cycloprop
43	146.5	8.0	389	2	B71865	cyclopropane-fatty
44	146.5	8.0	389	2	H64571	cyclopropane-fatty
45	144.5	7.9	241	2	C87607	hypothetical prote

ALIGNMENTS

RESULT 1

C96673 gamma-tocopherol methyltransferase [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: C96673

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiet, N.A.; Kaul, S.; White, O.

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; I

ansen, N.F.; Hughes, B.; Huizar, L.

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maltl, R.;

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: C96673

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-348 <STO>

A:Cross-references: GB:AE005173; NID:g8099780; PIDN:AAD38271.2; GSPDB:GNO0141

C:Genetics:

A:Gene: F13011.27

A:Map position: 1

Query Match 63.1%; Score 1154.5; DB 2; Length 348;

Best Local Similarity 69.4%; Pred. No. 36-87;

Matches 227; Conservative 28; Mismatches 61; Indels 11; Gaps 4

32 GPRSWAPIRASAAASE-----RGEIVLEQPKDKDKKLOKIAEFYDESSGLMWNIG 85

23 GSKSLIFRSPSSSSSSMTTGNVAANAATSTE--ALRKIAEYNTSLMEIING 80

86 DHNHGFEYDSDFVSLSDD--HRAAOIRMOIESLRFASVS--EERSKMPKSLVDVCGIGGS 142

81 DHNHGFEYDSSVQSLSDSHKKAQIRIMEESLRFAGYVDEEEKIKKKVVDVCGIGGS 140

143 SRYLAKFGATSVGITLSPVQAQANALAAAGLADKVSQVADALQOPSSDGFVLWS 202

141 SRYLAKFGATSVGITLSPVQAQANALAAAGLADKVSQVADALQOPSSDGFVLWS 200

203 MESGEHMPDKAKVGVGLARAAAGATITTYTWCNRDLPPEOSLHPEODLLKIKDAY 262

201 MESGEHMPDKAKVGVGLARAAAGATITTYTWCNRDLPPEOSLHPEODLLKIKDAY 260

263 LPWMCSTDDVYVVKLQSLQSDIRKSDMSRFVAPFPVAFVSAFTWKLSSLSGOKTK 322

261 LPWMCSTDDVYVVKLQSLQSDIRKSDMSRFVAPFPVAFVSAFTWKLSSLSGOKTK 320

323 GALAMPLEIEGKKDLIKFAITTCRKP 349

Db 321 GATMLPIMIEGKGVYKREITTCQKP 347

RESULT 2

gamma-tocopherol methyltransferase (imported) - Nostoc sp. (strain PCC 7120)
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Species: Nostoc sp.
 C.Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
 C.Accession: AE2031
 R.Kaneoko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A.Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
 A.Reference number: AB1807; MUID:21595285; PMID:11759840
 A.Accession: AE2031
 A.Status: Preliminary
 A.Molecule type: DNA
 A.Residues: 1-280 <KUR>
 A.Cross-references: GB:BA000019; PIDN:BA073502.1; PID:g17130893; GSPDB:GN00179
 A.Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr1803

Query Match 36.0%; Score 659; DB 2; Length 280;
 Best Local Similarity 49.3%; Pred. No. 1.3e-46;

Matches 138; Conservative 42; Mismatches 92; Indels 8; Gaps 3;

QY 65 LQKGAIEFYDESSGLMENIMGDMHGHGYDSSTVSLSDHRAAOIRMIQESLRFAVSSEE 124
 Db 5 LYQOIQOFPYDASSGIMEITWSEHMHGYYGADGT-EQKNRQAOIDLLEELTAVG--- 60
 QY 125 RSKMPKSIYDVCGGIGSSRYLAKFKGATSVGITLSPVOAQRANALAAAGLADVSFOY 184
 Db 61 --QTAENTLDYVCGGIGSSRYLAKFKGATSVGITLSPVOAQRANALAAAGLADVSFOY 118
 QY 185 ADALQOFPSPDGFPLVMSMEGEMHPDKAKFVGLARVAPGAIITITVWCHRDLPDEQ 244
 Db 119 ANAQAPFPDDNSFDLWMSLESSEHMPDKFLOECRYLAKFGKGLIWTWCHRP--TDKT 176
 QY 245 SLHPWEDLLKIKIDAVYLPAMCSTSDYVKLQSLSDIKSEDSMRFAVAPFVAVIRSA 304
 Db 177 PLTADERRHLQAIYDYVCLPYVSLPDYAIARECGFGEIKTADMSVAAPFVAVIRSA 236
 QY 305 FTWKGLSSLSGQRTIKGALAMPIMEGKRDILKFAITCRKP 349
 Db 237 FTPOAIFGLRAGWTTIGALISGLMRGREGILIRFGL 276

RESULT 3

hypothetical protein - Synechocystis sp. (strain PCC 6803)
 C:Species: Synechocystis sp.
 A:Variety: PCC 6803
 C.Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
 C.Accession: S76618
 R.Kaneoko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
 DNA Res. 3, 109-136, 1996
 A.Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis

S.
 A:Reference number: S74322; MUID:97061201; PMID:8905231
 A.Accession: S76618
 A.Status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-317 <KAN>
 A.Cross-references: EMBL:D64004; GB:AB001339; NID:g1001701; PIDN:BA010562.1; PID:d101121
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
 C:Superfamily: bioC homology
 F:95-199/Domain: bioC homology <BIOC>

Query Match 35.4%; Score 647.5; DB 2; Length 317;
 Best Local Similarity 47.4%; Pred. No. 1.3e-45;

Matches 135; Conservative 45; Mismatches 102; Indels 3; Gaps 3;

QY 65 LQKGAIEFYDESSGLMENIMGDMHGHGYDSSTVSLSDHRAAOIRMIQESLRFAVSSEE 124
 Db 34 LYEKIRNFYDSSGGLMEDVWGEHMHGYYPHOTYRI-DKROQOIDLILKELMAA-VPOH 91
 QY 125 RSKMPKSIYDVCGGIGSSRYLAKFKGATSVGITLSPVOAQRANALAAAGLADVSFOY 184
 Db 92 SAK-PKRIIDLDGCGIGSSRYLAKFKGATSVGITLSPVOAQRANALAAAGLADVSFOY 150
 QY 185 ADALQOFPSPDGFPLVMSMEGEMHPDKAKFVGLARVAPGAIITITVWCHRDLPDEQ 244
 Db 151 ANALDIPFASDSFDWMSLESSEHMPDKFLOECRYLAKFGKGLIWTWCHRPIDGNG 210
 QY 245 SLHPWEDLLKIKIDAVYLPAMCSTSDYVKLQSLSDIKSEDSMRFAVAPFVAVIRSA 304
 Db 211 PLTADERRHLQAIYDYVCLPYVSLPDYAIARECGFGEIKTADMSVAAPFVAVIRSA 270
 QY 305 FTWKGLSSLSGQRTIKGALAMPIMEGKRDILKFAITCRKP 349
 Db 271 FDPRLVLMALGQAGPKIINNALCLRLMKWCEGVLVRFGLTGIKP 315

RESULT 4

hypothetical protein al12121 (imported) - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp.

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C.Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
 C.Accession: AC2071
 R.Kaneoko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.;
 DNA Res. 8, 205-213, 2001

A.Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacter
 A.Reference number: AB1807; MUID:21595285; PMID:11759840
 A.Accession: AC2071
 A.Status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-330 <KUR>
 A.Cross-references: GB:BA000019; PIDN:BA073820.1; PID:g17131212; GSPDB:GN00179
 A.Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: al12121

Query Match 21.4%; Score 392.5; DB 2; Length 330;
 Best Local Similarity 34.7%; Pred. No. 1.3e-24;
 Matches 102; Conservative 50; Mismatches 101; Indels 41; Gaps 12;

QY 69 IAEFYDE--SSGLMENIMGDMHGHGYDS-----DSTVSLSDHRAAOIRMIQESLRFAVS 121
 Db 35 VANSYDQWTEGDIIEFYTGSHIHGHGSPORRDFVYAKSD-----FVHEWVWNGGL 87
 QY 122 SEERSKWP--KSIYDVCGGIGSSRYLAKFKGATSVGITLSPVOAQRANALAAAGLADK 179
 Db 88 ----DKLPETITLDVCGGIGSSRYLAKFKGATSVGITLSPVOAQRANALAAAGLADK 141
 QY 180 VSRQVADALQOFPSPDGFPLVMSMEGEMHPDKAKFVGLARVAPGAIITITVWCHRD 239
 Db 142 -QFLVADAMALSPDDNSFDWMSLESSEHMPDKFLOECRYLAKFGKGLIWTWCHRPD 199
 QY 240 GPEQDS-LHPWEDLLKIKIDAVYLPAMCSTSDYVKLQSLSDIKSEDSMRFAVAPF 296
 Db 200 --DRQKPLNFWKPEVWQQLDQSHPAFSSIEGFSLLATGATGVECEVITADWTOTLPS 257
 QY 297 WPAVIRSAFTWGL---SSL---LSSGQRTIKGALAMPIMEGKRDILKFAITCRKP 343
 Db 258 WLDISI-----WQGIYVPEGLVRFGLSGFKSLAREVPTLLMLRLARETGICRFEGM 306

RESULT 5

hypothetical protein - Synechocystis sp. (strain PCC 6803)
 C:Species: Synechocystis sp.

RESULT 8
T44579
C5-O-methyltransferase [imported] - Streptomyces avermitilis
C:Species: Streptomyces avermitilis
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C:Accession: T44579
R:Ikeda, H.; Nonomiya, T.; Usami, M.; Ohta, T.; Omura, S.
Proc. Natl. Acad. Sci. U.S.A. 96, 9509-9514, 1999
A:Title: Organization of the biosynthetic gene cluster for the polyketide anthelmintic m
A:Reference number: 222796; MUID:99380548; PMID:10449723
A:Accession: T44579
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-283 <IKE>
A:Cross-references: EMBL:AB032524; NID:g5921164; PIDN:BAAB4602.1; PID:g5921167
C:Genetics:
A:Note: aved

Query Match 13.0%; Score 238; DB 2; Length 283;
Best Local Similarity 26.3%; Pred. No. 5.4e-12;
Matches 80; Conservative 51; Mismatches 135; Indels 38; Gaps 10;

OY 57 PKDDKKRLQKIAEFYDESSGLMWNHGFYSDSTVSLSDHRAAQIRMIQESL 116
DB 5 PEQTPSSLEVG--DYDRLTDLNMRALGNTHLGYWPOPDGS--SPGKAA-----DRL 55
OY 117 RFASVSEERSKMPKSIYVCGGIGSSRYLAKKRGATSVGTTLSFYQAQRANALAAAGL 176
DB 56 TDLLGKRLGRTGRVRLVCGGSGKPAVRLASAPVDVGVYSEVQGLATLAKOSH 115
OY 177 ADKVSFOVADALQOFPSPGDFDLVMSMSEGHMPDKAFVGEELARVAAPGAIITIVMCH 236
DB 116 ADKVVFTPADHMLPFPDGSFDAMALCLLHNPSPAVYIEIARVLRPGRLAVTDVAL 175
OY 237 RDLGPDEQSLHPWEODLLKIKCDAYVLPAMCSTSDYVVLQSLSDIKSED--NSRFVAP 295
DB 176 RAFGTGKMKRGECTSOLLA-----VPALVHIDEVACIMADAGLELHEDLDIGQVYGP 228
OY 296 FMPAV-----IRSAFTWKGLSLSSGOKTIRKALAMPIMIEYKKDLIKFAIT 345
DB 229 SFALRDHVNHEHDEYAAAF--GIG--VAEMRKVVAOCTTLP-----WTFD-IGYVVL 277
OY 346 CRKP 349
DB 278 ARRP 281

RESULT 9
T04138
sterol 24-C-methyltransferase (EC 2.1.1.41) ESM1, endosperm - maize
N:Alternate names: C-24 sterol methyltransferase
C:Species: Zea mays (maize)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 03-Jun-2002
C:Accession: T04138
R:Grebencok, R.J.; Galbraith, D.W.; Dellapenna, D.
Plant Mol. Biol. 34, 891-896, 1997
A:Title: Characterization of zea-mays endosperm C-24 sterol methyltransferase - one of 2
A:Reference number: 209668; MUID:97435974; PMID:9290641
A:Accession: T04138
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-344 <GRE>
A:Cross-references: EMBL:U79669; NID:g1899059; PIDN:AA870886.1; PID:g1899060
A:Experimental source: endosperm
C:Genetics:
A:Gene: ESM1
C:Function:
A:Description: methyltransferase
C:Superfamily: 24-sterol C-methyltransferase; bioc homology
C:Keywords: methyltransferase; S-adenosylmethionine
F:101-205/Domain: bioc homology <BIOC>

Query Match 12.4%; Score 227; DB 2; Length 344;

Best Local Similarity 24.6%; Pred. No. 5.6e-11;
Matches 82; Conservative 47; Mismatches 109; Indels 96; Gaps 12;

OY 69 IAEFYDESSGLMWNHGFYSDSTVSLSDHRAAQIRMIQESLRFASVSEERSK 128
DB 51 VNKYDALTATSEYEGMGCSFHFARHNGESLRESIKRHEHFLAOLGL-----KP 100
OY 129 PKSTVDCCGIGGSSRYLAKKRGATSV--GTTLSFYQAQRANALAAAGLADKVSFOVADA 187
DB 101 GKNVLDVCGGIGGPRLETA--RFSSTSVGLNNNEQITRGKELNLAISGTCDFVKA 159
OY 188 LQOFPDQOFLVMSMSEGHMPDKAFVGEELARVAAPGAIITIVMCH-----H 236
DB 160 MKMFPDDWTFPAVVAIEATCHAPDPVCGCYKEIRVRLKGGCFAYEMCITDHYDPNMT 219
OY 237 R-----DLG---PDEQSLHP-----WEODLLKIKCDAYVLPAMCS----- 268
DB 220 KRKDEIELGNLPDIRSTROCLQAVKDAGEVWDKDLAEDSLPMLYPLDPSRFSLS 279
OY 269 -----TSDYKLLQSLSDIKSEDSRFVAPRPVIRSAFTWKGLSLSSGO 318
DB 280 FRLTSVGRATRTWVKALEYGLAPQGSERVSNFL-----E 315
OY 319 KTIKALAMPIMIEYKKDL--IKFAITCRKP 349
DB 316 KAAEC-----LVEGKKEIFTPMYFFLV--RKP 341

RESULT 10
JC6531
avermetrin B 5-O-methyltransferase (EC 2.1.1.-) - Streptomyces "avermitilis"
C:Species: Streptomyces "avermitilis"
C:Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 16-Jul-1999
C:Accession: JC6531
R:Ikeda, H.; Wang, L.R.; Ohta, T.; Inokoshi, J.; Omura, S.
Gene 206, 175-180, 1998
A:Title: Cloning of the gene encoding avermetrin B 5-O-methyltransferase in avern
A:Reference number: JC6531; MUID:98137789; PMID:9469930
A:Accession: JC6531
A:Molecule type: DNA
A:Residues: 1-283 <IKE>
A:Experimental source: wild type K139
C:Comment: This enzyme catalyzes the methylation of the hydroxyl group at the C5
C:Genetics:
A:Gene: aved
C:Superfamily: bioc homology
C:Keywords: methyltransferase
F:68-172/Domain: bioc homology <BIOC>

Query Match 12.3%; Score 226; DB 2; Length 283;
Best Local Similarity 26.0%; Pred. No. 5.3e-11;
Matches 79; Conservative 50; Mismatches 137; Indels 38; Gaps 10;

OY 57 PKDDKKRLQKIAEFYDESSGLMWNHGFYSDSTVSLSDHRAAQIRMIQESL 116
DB 5 PEQTPSSLEVG--DYDRLTDLNMRALGNTHLGYWPOPDGS--SPGKAA-----DRL 55
OY 117 RFASVSEERSKMPKSIYVCGGIGSSRYLAKKRGATSVGTTLSFYQAQRANALAAAGL 176
DB 56 TDLLGKRLGRTGRVRLVCGGSGKPAVRLASAPVDVGVYSEVQGLATLAKOSH 115
OY 177 ADKVSFOVADALQOFPSPGDFDLVMSMSEGHMPDKAFVGEELARVAAPGAIITIVMCH 236
DB 116 ADKVVFTPADHMLPFPDGSFDAMALCLLHNPSPAVYIEIARVLRPGRLAVTDVAL 175
OY 237 RDLGPDEQSLHPWEODLLKIKCDAYVLPAMCSTSDYVVLQSLSDIKSED--NSRFVAP 295
DB 176 RAFGTGKMKRGECTSOLLA-----VPALVHIDEVACIMADAGLELHEDLDIGQVYGP 228
OY 296 FMPAV-----IRSAFTWKGLSLSSGOKTIRKALAMPIMIEYKKDLIKFAIT 345
DB 229 SFALRDHVNHEHDEYAAAF--GIG--VAEMRKVVAOCTTLP-----WTFD-IGYVVL 277

C; Superfamily: 24-sterol C-methyltransferase; bioc homology

QY 249 WEQDLKKICDAYYLEPANCSTSDYVKLLQSLSLQDIKSED 288

249 WEQDLKKICDAYLLPWCST
QY

QY 249 WEEDLKKICDAYYLPWCSTSDYVKLLQSLSLQDIKSED 288

QY 249 WEQDLKIKIDAYLPAMCSISDYVKLIQSLSLDIKSED 288

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 17:36:22 ; Search time 11 Seconds
(without alignments)
1319.701 Million cell updates/sec

Title: US-09-857-613A-28

Perfect score: 1830

Sequence: 1 MATVVRIPRTSCIHIFRS.....IEGKKDLIFAITCKRPE 350

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : Swissprot_40:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	221.5	12.1	378	1 ERG6_SCHPO	014321 schizosacch
2	203.5	11.1	382	1 ERG6_YEAST	P25008 saccharomyc
3	189	10.3	376	1 ERG6_CANAL	074198 candida alb
4	174	9.5	394	1 YLP3_PSEPU	P31049 pseudomonas
5	156	8.5	274	1 YEO3_MYCTU	P71673 mycobacteri
6	152	8.3	381	1 CPA_ECOLI	P30010 escherichia
7	143	7.8	251	1 BIOC_ERWHE	006898 erwinda her
8	141	7.7	251	1 Y095_HARTN	057060 haemophilus
9	137	7.5	238	1 UBIG_PASMO	087025 talstonia s
10	134.5	7.3	242	1 UBIG_PASMO	09cm16 pasteurella
11	131.5	7.2	289	1 UBIG_NEIMB	09jx17 neisseria m
12	131.5	7.2	289	1 UBIG_RICCN	092h07 rickettsia
13	130	7.1	274	1 YEO3_MYCTU	P71673 mycobacteri
14	129.5	7.1	238	1 UBIG_NEIMA	09jxw6 neisseria m
15	128	7.0	232	1 UBIG_PSENE	09h263 pseudomonas
16	125.5	6.9	244	1 YXBB_BACSD	P46326 bacillus su
17	124.5	6.8	240	1 UBIG_ECO57	08xex2 escherichia
18	124.5	6.8	252	1 UBIG_CACOR	09a9x1 caulobacter
19	123.5	6.7	240	1 UBIG_ECOLI	P17993 escherichia
20	121.5	6.6	258	1 TAM_DEIRA	09xrx3 deinococcus
21	118.5	6.5	248	1 UBIG_RHIME	09amk1 rhizobium m
22	117	6.4	242	1 UBIG_YERPE	08zgr6 yersinia pe
23	117	6.4	252	1 UBIG_RICPR	09zct9 rickettsia
24	117	6.4	257	1 YIG4_YEAST	P40516 saccharomyc
25	116	6.3	203	1 PMTA_RHOSH	005197 rhodobacter
26	115.5	6.3	242	1 UBIG_SALTI	08x560 salmonella
27	115.5	6.3	242	1 UBIG_SALTY	P37431 salmonella
28	115.5	6.3	269	1 RUMA_ECOLI	P36999 escherichia
29	115	6.3	220	1 RUMA_VIBS	P55136 vibrio sp.
30	113.5	6.2	417	1 YAT1_SYNP6	P08442 synecococc
31	113	6.2	212	1 Y829_SYNP3	055423 synecococc
32	111	6.1	246	1 UBIG_XYLFA	09pam5 xyella fas
33	110	6.0	245	1 UBIG_VIBCH	09ksj9 vibrio chol

34	110	6.0	296	1 PRMA_HAEIN	P44402 haemophilus
35	110	6.0	322	1 COO3_ARATH	049354 a hexapreny
36	107.5	5.9	248	1 UBIG_BRUME	08y198 brucella me
37	107.5	5.9	261	1 TAM_MYCTU	033698 mycobacteri
38	105	5.7	224	1 BCMH_RHOCA	P26236 rhodobacter
39	105	5.7	348	1 HMT1_YEAST	P38074 saccharomyc
40	104.5	5.7	249	1 UBIG_RHILLO	098987 rhizobium l
41	104	5.7	248	1 UBIE_RICCN	09zgt5 rickettsia
42	102.5	5.6	255	1 TAM_BRUME	08y191 brucella me
43	101	5.5	252	1 GRCH_LACIA	P49016 lactococcus
44	100.5	5.5	251	1 BIOC_ECOLI	P12999 escherichia
45	100.5	5.5	286	1 COO3_RAT	063159 r hexapreny

ALIGNMENTS

RESULT 1
ERG6_SCHPO STANDARD: PRT: 378 AA.
ID ERG6_SCHPO
AC 014321: P78782;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable delta(24)-sterol C-methyltransferase (EC 2.1.1.41).
GN SPEC169.05.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
CX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21848401; PubMed=11859360;
RX Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
Brooks K., Brown D., Brown S., Chillingworth J., Churcher C.M.,
Collins M., Connor R., Davis P., Davis P., Feltwell T., Fraser A.,
Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeli C.,
Olliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
Skellton J., Simmonds M., Squares R., Squares S., Stevens K.,
Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
Widjaja J., Woilckaert G., Aert R., Robben J., Grymoprez B.,
Weldjens J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
Eger P., Zimmermann M., Wedler H., Wambutt R., Purnelle B.,
Goffeau A., Cadieu E., Dreano S., Gloux S., Lejaune V., Motier S.,
Gallbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
Lucas M., Rochet M., Galliard C., Tallada V.A., Garçon A., Thode G.,
Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
Domínguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
Shpakowski G.V., Ussery D., Barrett B.G., Nurse P.;
"The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
RN [2]
RP SEQUENCE OF 55-378 FROM N.A.
RC STRAIN=PR745;
RX MEDLINE=98162722; PubMed=9501991;
RA Yoshida S., Kato K., Nakai K., Okayama H., Nojima H.;
"Identification of open reading frames in Schizosaccharomyces pombe
cDNAs.";
RT DNA Res. 4:363-369(1997).
CC -I- FUNCTION: METHYLTRANSFERASE REQUIRED FOR ERGOSTEROL SYNTHESIS (BY
SIMILIARITY).
CC -I- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 5-alpha-cholest-
8,24-dien-3-beta-ol -> S-adenosyl-L-homocysteine + 24-methylene-5-

```

CC      alpha-cholest-8-en-3-beta-ol.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL; Z99759; CAB16897.1; -
DR      EMBL; D89131; BAA13793.1; -
DR      InterPro: IPR001601; Methyltransf.
DR      InterPro: IPR000051; SAM bind.
DR      InterPro: IPR004033; UbiE/COQ5_Metrf.
DR      Pfam; PF01209; UbiE_methyltran; 1.
KM      Sterol biosynthesis; Transferase; Methyltransferase.
FT      CONFLICT 55 63 NHEMSEED -> SVPGDPLES (IN REF. 2).
SQ      SEQUENCE 378 AA; 42867 MW; FA4D3D82A1CE03D6 CRC64;

Query Match          12.1%; Score 221.5; DB 1; Length 378;
Best Local Similarity 22.5%; Pred. No. 3.4e-11;
Matches 63; Conservative 46; Mismatches 124; Indels 47; Gaps 5;

OY      39 IRASASSEGEIVLEOKPKKDKKRLQK-----GIAEF 72
DB      19 LHGKAAREKKTGLAATASKNVDQSKRLQEFEFMDRNHNESEEDRARIDGKSVNXY 78
OY      73 YDESGLEWNTWGDHMHG-PYDSDS---TVSLSPHRAQIMIGESLRFASVSEBSKRW 128
DB      79 YDLADLYEYGSQSFHSFYKGAFAQSIARHEHYLAIRMGIRKPSR----- 127
OY      129 PKSIYVCGGIGGSSRYLAKKFGATSVGTTLSPVQANALAAOGLADKVSFOVADAL 188
DB      128 ---YLDVCGVGCPAREITEFTGICNLVGLNNNDYQISRCNNAYAVRNLDKKYFKGDPW 184
OY      189 QOPFSDGQFDLVMSGESGEMPDKAFYVELARVAAPGAIITIVTCHNDLGPDSQSLHP 248
DB      185 HHPFEDNPFDYVYAEATFVHAPSLGEGVGEIRVLKPGCVFVYEMVMSD---DYDSISP 241
OY      249 WEQDLKRICDAVYILPAMKSTSDYKLLQSLSLQDIKSD 288
DB      242 KHREIAYNIYGVGDIPIQVNRKCDAVEAIKKVGNLLEED 281

RESULT 2
ERG6_YEAST
ID      ERG6_YEAST          STANDARD:          PRT:          382 AA.
AC      P25087;
DT      01-MAY-1992 (Rel. 22, Created)
DT      01-OCT-1994 (Rel. 30, Last sequence update)
DT      01-OCT-1996 (Rel. 34, Last annotation update)
DE      Delta(24)-sterol C-methyltransferase (EC 2.1.1.41).
GN      ERG6 OR SED6 OR ISE1 OR LIS1 OR YML008C OR YM9571.10C.
OS      Saccharomycetes cerevisiae (Baker's yeast).
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC      Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX      NCBI_TaxID=4932;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=94262330; PubMed=8203167;
RA      Hardwick K.G., Pelham H.R.B.;
RT      "SED6 is identical to ERG6, and encodes a putative methyltransferase
RT      required for ergosterol synthesis.";
RL      Yeast 10:265-269(1994).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=94312403; PubMed=8038180;
RA      Wellhinda A.A., Pavis A.D., Trumbly R.J.;
RT      "Mutations in ... (ERG6) gene confer increased sodium and lithium
RT      uptake in Saccharomycetes cerevisiae.";
RL      Biochim. Biophys. Acta 1193:107-117(1994).
RN      [3]

```

```

RP      SEQUENCE FROM N.A.
RC      STRAIN=S288c / AB972;
RA      Gentles S., Bowman S., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL      Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN      [4]
RP      SEQUENCE OF 1-257 FROM N.A.
RX      MEDLINE=91285426; PubMed=2060792;
RA      Hussain M., Lenard J.;
RT      "Characterization of PDR4, a Saccharomyces cerevisiae gene that
RT      confers pleiotropic drug resistance in high-copy number: identity
RT      with YAP1, encoding a transcriptional activator.";
RL      Gene 101:149-152(1991).
RN      [5]
RP      ACETYLATION.
RA      Garrels J.I., Fletcher B., Kobayashi R., Latter G.I., Schwender B.,
RA      Volpe T., Warner J.R., McLaughlin C.S.;
RL      Submitted (SEP-1994) to the SWISS-PROT data bank.
CC      -1- FUNCTION: METHYLTRANSFERASE REQUIRED FOR ERGOSTEROL SYNTHESIS.
CC      -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 5-alpha-cholest-
CC      8,24-dien-3-beta-ol -> S-adenosyl-L-homocysteine + 24-methylene-5-
CC      alpha-cholest-8-en-3-beta-ol.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL; X74249; CAA52308.1; -
DR      EMBL; S72460; AAB31378.1; -
DR      EMBL; Z49810; CAA89944.1; -
DR      EMBL; X53830; CAA37826.1; -
DR      PIR; PE0402; PE0402.
DR      PIR; S17001; S17001.
DR      PIR; S35982; S35982.
DR      PIR; S42003; S42003.
DR      SWISS-2DPAGE; P25087; YEAST.
DR      SGD; S0004467; ERG6.
DR      InterPro: IPR001601; Methyltransf.
DR      InterPro: IPR000051; SAM bind.
KM      Sterol biosynthesis; Transferase; Methyltransferase; Acetylation.
FT      INIT MET 0 0
FT      MOD. RES 1 1 ACETYLATION.
FT      CONFLICT 379 379 E -> EE (IN REF. 2).
SQ      SEQUENCE 382 AA; 43299 MW; D76BAE9E3D9CD71B CRC64;

Query Match          11.1%; Score 203.5; DB 1; Length 382;
Best Local Similarity 21.8%; Pred. No. 1.1e-09;
Matches 78; Conservative 58; Mismatches 149; Indels 73; Gaps 12;

OY      42 SAASSEGEIVLEOKPKKDKKRLQKGLAEFDESGGLMENT-----WGDHMHG-F 92
DB      35 NSAQREAVQKYLIRNMDGRDKDAEERRL-EDYNEATHSYNVNTDYFYEYGSQSFHSRFS 93
OY      93 YDSDS---TVSLSDHRAAQMIGESLRFASVSEBSKWKPSIVDVCGGIGSSRYLAKK 149
DB      94 YKGESEFASINHEHYLAIRMGIRK-----DVLVDGCGVGCPAREIARF 139
OY      150 FGATSVGTTLSPVQANALAAOGLADKVSFOVADALQOPFSDGQFDLVMSGESGEMH 209
DB      140 TGCNVIIGLNNDYQIAKAKYKYNLSQDMDFVKGDPKMDFEENFDKVAIEATCHA 199
OY      210 PDKAFVGEIARVAAPGAIITIVTCHNDLGPDSQSLHPWEDLKKIKTCDAVY-----L 263
DB      200 PKLEGVSEITVKVLPKPGTFAYEVWMTDKYDENPHE-----RKI--AYEELGQCI 250
OY      264 PFWCSTSDYVKLLQSLSLQDIKSD-----NSRFVAFWPVAV-----IRSAF 305
DB      251 PMMFHVDAARAKLAKNGCEVLVSEDLADNDDEIPIYVLTGEMKAYVQULANLATPFRSY 310
OY      306 TWKGLSS-----LLSGQKTIKAL---AMPLEIMIEGKKDLIRFAITTCRKE 350

```

Db 311 LGRFTTAMTYMEKLGLEAPGSKVEYALLENAAVGLVAGCKSLFTPMHLPVARKPE 368

RESULT 3

ERG6_CANAL STANDARD; PRT; 376 AA.

AC 074198;

DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DE Delta(24)-sterol C-methyltransferase (EC 2.1.1.41).

GN ERG6

OS Candida albicans (Yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Mitosporic Saccharomycetales; Candida.

NCBI_TaxID=5476;

RP SEQUENCE FROM N.A.

RX MEDLINE-98253976; PubMed-9593144;

RA Jensen-Pergakes K.L., Kennedy M.A., Lees N.D., Barbuch R., Koegele C.,

RA Bard M.;

RT "Sequencing, disruption, and characterization of the Candida albicans

sterol methyltransferase (ERG6) gene: drug susceptibility studies in

erg6 mutants." Agents Chemother. 42:1160-1167(1998).

RL Anticrib. Agents Chemother. 42:1160-1167(1998).

CC -1- FUNCTION: METHYLTRANSFERASE REQUIRED FOR ERGOSTEROL SYNTHESIS.

CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 5-alpha-cholest-

8,24-dien-3-beta-ol -> S-adenosyl-L-homocysteine + 24-methylene-5-

alpha-cholest-8-en-3-beta-ol.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

CC EMBL; AF031941; AAC26626.1;

DR Interpro: IPR001601; Methyltransf.

DR Interpro: IPR000051; SAM_bind.

KW Sterol biosynthesis; Transferase; Methyltransferase.

SQ SOURCE 376 AA; 43085 MW; 369094aFCDDEDX0 CRC64;

Query Match 10.3%; Score 189; DB 1; Length 376;

Best Local Similarity 21.6%; Pred. No. 1.8e-08;

Matches 74; Conservative 55; Mismatches 133; Indels 80; Gaps 13;

Y 59 KDKKKLQKGIAE---FYDESSGLMENINGDHNHG-FYDSDS---TVSLSDHRAAOIR 110

Db 59 KDKKKLQKGIAE---FYDESSGLMENINGDHNHG-FYDSDS---TVSLSDHRAAOIR 110

Y 111 MIDESLRFASVSEERSKWPISYDVCGGSSRYLAKKFGATSVGTLISVVOQRANAL 170

Db 111 MIDESLRFASVSEERSKWPISYDVCGGSSRYLAKKFGATSVGTLISVVOQRANAL 170

Y 119 NLENENK-----VDVCGGCGVGGPGEITRFDCETVGLANNNOYQIERANHY 164

Db 119 NLENENK-----VDVCGGCGVGGPGEITRFDCETVGLANNNOYQIERANHY 164

Y 171 AAAAGLADKVSFOYADALQOPFSDGQFDLVMSGEGHMPDKAFVGLAFAVAAPGATII 230

Db 171 AAAAGLADKVSFOYADALQOPFSDGQFDLVMSGEGHMPDKAFVGLAFAVAAPGATII 230

Y 165 AKKTHLDHKLISYVKGDMQDFEPESDAVAIATVAHAPLDEGVSEIYVLAAPGIFG 224

Db 165 AKKTHLDHKLISYVKGDMQDFEPESDAVAIATVAHAPLDEGVSEIYVLAAPGIFG 224

Y 231 IVTWC-----HRDLG-----PDEOSLH--PWEODLKKICDA- 260

Db 231 IVTWC-----HRDLG-----PDEOSLH--PWEODLKKICDA- 260

Y 225 VYEVMTDKYDETNEHRKIAVGLGVGDIPIKMSRKVAEQALNKGVEIYQKDLAVD 284

Db 225 VYEVMTDKYDETNEHRKIAVGLGVGDIPIKMSRKVAEQALNKGVEIYQKDLAVD 284

Y 261 ----YLP-----AMCST--SDYVKLQSLSLQDIKSEDSFVAPFPAVAFRSAPFTWGL 310

Db 261 ----YLP-----AMCST--SDYVKLQSLSLQDIKSEDSFVAPFPAVAFRSAPFTWGL 310

Y 311 SSLSSGQTKIKAL---APLMLEGGYKKDLIKFAIITCRAP 349

Db 311 SSLSSGQTKIKAL---APLMLEGGYKKDLIKFAIITCRAP 349

Y 331 GLAPKSGKQVTHALDEDAVAVLVGCGKLEFTPMHLPVARKPE 371

Db 331 GLAPKSGKQVTHALDEDAVAVLVGCGKLEFTPMHLPVARKPE 371

RESULT 4

YLP3_PSEPU STANDARD; PRT; 394 AA.

AC P31049;

DT 01-JUL-1993 (Rel. 26, Created)

DT 01-JUL-1993 (Rel. 26, Last sequence update)

DE Hypothetical 44.7 kDa protein in LpD-3.5' region (ORF3).

OS Pseudomonas putida.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas.

NCBI_TaxID=303;

RP SEQUENCE FROM N.A.

RX STRAIN-PG2;

RA Lorenz D., Sokatch J.R.;

RL Submitted (MUG-1992) to the EMBL/GenBank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

CC EMBL; X55704; CAA39234.1;

DR Interpro: IPR003333; CMAS.

DR Interpro: IPR001601; Methyltransf.

DR Interpro: IPR000051; SAM_bind.

KW Hypothetical protein.

SQ SEQUENCE 394 AA; 44656 MW; 8CD61FE9495302AF CRC64;

Query Match 9.5%; Score 174; DB 1; Length 394;

Best Local Similarity 30.5%; Pred. No. 3.5e-07;

Matches 51; Conservative 30; Mismatches 58; Indels 28; Gaps 7;

Y 54 EOKPKK--DDKKLQKGIAEFYDESSGLMENINGDHNHG---FYDSSTVLSL----- 102

Db 54 EOKPKK--DDKKLQKGIAEFYDESSGLMENINGDHNHG---FYDSSTVLSL----- 102

Y 90 EQPERSHDKRTAEALISYHYDVSNATYQ-LMDQDMAYGAYREPDNTLDQAQDKF 148

Db 90 EQPERSHDKRTAEALISYHYDVSNATYQ-LMDQDMAYGAYREPDNTLDQAQDKF 148

Y 103 DHPAAQIRMIQESLRFASVSEERSKWPISYDVCGGSSRYLAKKFGATSVGTLSPV 162

Db 103 DHPAAQIRMIQESLRFASVSEERSKWPISYDVCGGSSRYLAKKFGATSVGTLSPV 162

Y 149 DHLCKRLKLNAGDY-----LIDVCGGCGMGLAFARREYDAVAFGTLISKE 193

Db 149 DHLCKRLKLNAGDY-----LIDVCGGCGMGLAFARREYDAVAFGTLISKE 193

Y 163 QAORANLMAAGLADKVSFOYADALQOPFSDGQFDLVMSGEGHMPDKAFVGLAFAVA 209

Db 163 QAORANLMAAGLADKVSFOYADALQOPFSDGQFDLVMSGEGHMPDKAFVGLAFAVA 209

Y 194 QLKGRQRYKAEGLTDKVDQILIDYRDL-ODGRFDKVVSYGMFEBV 239

Db 194 QLKGRQRYKAEGLTDKVDQILIDYRDL-ODGRFDKVVSYGMFEBV 239

RESULT 5

YEO5_MYCTU STANDARD; PRT; 274 AA.

AC P71673;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DE Hypothetical protein RY1405C precursor.

GN RY1405C OR MY1449 OR MYCY21BA.22C.

OS Mycobacterium tuberculosis.

OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;

OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.

NCBI_TaxID=1773;

RP SEQUENCE FROM N.A.

RX STRAIN-H37RV;

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,

RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,

RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,

RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sultison J.E., Taylor K., Whitehead S., Barrell B.G.:
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CDC 1551 / Oshkosh:
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.:
 RA "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: STRONG, TO M.TUBERCULOSIS RV1403C AND SOME, TO
 CC R.SPAREOIDES PMTA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: Z80108; CAB02184.1; -
 DR EMBL: AE007016; AAK45714.1; -
 DR TIGR: MT1449; -
 DR TubercuList: RV1405C; -
 DR InterPro: IPR000051; SAM_bind.
 KM Hypothetical protein; Signal; Complete proteome.
 FT SIGNAL 1 30
 FT CHAIN 31 274
 FT SEQUENCE 274 AA; 29262 MW; 3CF2932FE9E3531CA CRC64;
 SQ
 Query Match 8.5%; Score 156; DB 1; Length 274;
 Best Local Similarity 36.1%; Pred. No. 7e-06;
 Matches 43; Conservative 16; Mismatches 56; Indels 4; Gaps 3;
 QY 132 IVDCGIGGSSRYLAKFGATSVGITLSPVOAQRANALAAAGLADVSFOVADALQOP 191
 Db 53 VLDVAAGSGNIS-LPAATGATVISTDTLPPLQLRSGARAQOGIT-LQYQENAAQLP 109
 QY 192 FSDGQFDLWMSGEHMPDKAFVGEIARVAAPGAIITVTW-CHRDLPDEQSLHPW 249
 Db 110 FADDEFPTVISAIGVFAPDQAADDELIVRCRPGGTIGVISWCEGFGHMLATIRPY 168
 RESULT 6
 CFA_ECOLI STANDARD; PRT; 381 AA.
 ID CFA_ECOLI
 AC P30010;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cyclopropane-fatty-acyl-phospholipid synthase (EC 2.1.1.79)
 DE (Cyclopropane fatty acid synthase) (CFA synthase).
 GN CFA OR CDF A OR B1661.
 OS Escherichia coli.
 OC Bacteria: Proteobacteria: gamma subdivision: Enterobacteriaceae:
 OC Escherichia
 OC NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A., SEQUENCE OF 1-8, AND CHARACTERIZATION.
 RC STRAIN-K12;
 RX MEDLINE-93075691; PubMed-1445840;
 RA Wang A.-Y., Grogan D.W., Cronan J.E. Jr.:
 RT "Cyclopropane fatty acid synthase of Escherichia coli: deduced amino
 RT acid sequence, purification, and studies of the enzyme active site.";

RL Biochemistry 31:11020-11028(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / RR28;
 RA Eberhardt S.M.R., Richter G., Gimbel W., Werner T., Bacher A.:
 RL Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE-97426617; PubMed-9278503;
 RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Colloido-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.D.,
 RA Mau B., Shao Y.:
 RA "The complete genome sequence of Escherichia coli K-12."
 RT Science 277:1453-1474(1997).
 RL [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE-97251357; PubMed-9097039;
 RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
 RA Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
 RA Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,
 RA Motomura K., Nakade S., Nakamura Y., Nashimoto H., Nishio Y.,
 RA Oshima T., Saito N., Sempel G., Seki Y., Sivasubaram S.,
 RA Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
 RA Yamamoto Y., Horinchi T.:
 RA "A 570-kb DNA sequence of the Escherichia coli K-12 genome
 RT corresponding to the 28.0-40.1 min region on the linkage map."
 RL DNA Res. 3:363-377(1996).
 CC -1- FUNCTION: TRANSFERS A METHYLENE GROUP FROM S-ADENOSYL-L-METHIONINE
 CC TO THE CIS DOUBLE BOND OF AN UNSATURATED FATTY ACID CHAIN
 CC RESULTING IN THE REPLACEMENT OF THE DOUBLE BOND WITH A METHYLENE
 CC BRIDGE.
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + phospholipid
 CC olefinic fatty acid -> S-adenosyl-L-homocysteine + phospholipid
 CC cyclopropane fatty acid.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: M98330; AAA23562.1; -
 DR EMBL: X69109; -; NOT_ANNOTATED_CDS.
 DR EMBL: AE000261; AAC74733.1; -
 DR EMBL: D90809; BAA15428.1; -
 DR EMBL: D90810; BAA15437.1; -
 DR PIR: A44292; A44292.
 DR Ecogene; EG11531; Cfa.
 DR InterPro: IPR003333; CFA.
 DR InterPro: IPR001601; Methyltransf.
 DR InterPro: IPR000051; SAM_bind.
 DR Pfam: PF02353; CMA5; 1.
 KM Transferase; Methyltransferase; Lipid synthesis; Complete proteome.
 FT INIT_MET 0
 FT CONFLICT 1 1 S -> R (IN REF. 2).
 FT CONFLICT 7 7 E -> G (IN REF. 2).
 FT CONFLICT 24 24 S -> N (IN REF. 2).
 FT CONFLICT 38 38 I -> T (IN REF. 2).
 FT SEQUENCE 381 AA; 43777 MW; 1F07B220C7E08ADF CRC64;
 SQ
 Query Match 8.3%; Score 152; DB 1; Length 381;
 Best Local Similarity 24.1%; Pred. No. 2.3e-05;
 Matches 61; Conservative 40; Mismatches 94; Indels 58; Gaps 12;
 QY 71 EYDESSGLMNINGDHHHG---FYDSSTVYSLSDHRAQIRHIOESLPRASVSEERSK 127
 Db 117 EHYDGLNDFSRMDPFWQYSCAYWKAD---NLESQAQAKLMKICERLD-----K 165


```

UBIG_RALSO
ID UBIG_RALSO STANDARD; PRT; 238 AA.
AC 08Y025;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 3-demethylubiquinone-9-3-methyltransferase (EC 2.1.1.64) (3,4-
DE dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHNB
DE methyltransferase).
GN UBIG OR RSC0898 OR RS04517.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Catolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Siegler P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
CC -1- CARBOLYTIC ACTIVITY: S-adenosyl-L-methionine + 3-
CC -1- demethylubiquinone-9 = S-adenosyl-L-homocysteine + ubiquinone-9.
CC -1- PATHWAY: Ubiquinone biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE UBIG/COQ3 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AL646061; CAD14600.1; ALT INT.
CC InterPro: IPR001601; Methyltransf.
CC InterPro: IPR000051; SAM_bind.
CC InterPro: IPR004033; UblE/COQ5_Metrf.
CC Pfam: PF01209; UblE_methyltran; 1.
CC Ubiquinone biosynthesis; Transferase; Methyltransferase;
CC Complete proteome.
CC KW
CC SO SEQUENCE 238 AA; 25894 MW; 18114F00BF315808 CRC64;

Query Match 7.5%; Score 137; DB 1; Length 238;
Best Local Similarity 28.6%; Pred. NO. 0.00023;
Matches 57; Conservative 26; Mismatches 82; Indels 34; Gaps 8;
OY 90 HGYDSDSTVS-LSDRHQIRMIQESLRFASVSEERSKPKSIYVGGCIGSSRYLAK 148
DB 21 HRMDNDSEKPLHEINPLMDWISTAPLAG-----KRVVDGCGGGLSSMAR 71
OY 149 KEGATSVGITLSPVOAORANALAAOGLADKVSFOVADAL--COOPSDGOFDLMWSMESG 206
DB 72 A-GANVKGIDLSRKALRVADLHSLGAVADVDEIAEAALANEP---GSFVDVTOMEM 127
OY 207 EHMFDKAKFYGELARVAARAGATIIITWCHRDGPRDEQSLHPEQDLKICAYY----- 262
DB 128 EHVDPDASVVRACATLVKPGHVFST-----IHRNAKAYLLAVIAEYVLNM 175
OY 263 LPAMCSTDPVKLLQSLSL 281
DB 176 LPR--GTHDYAKTIRPSEL 192

```

```

RESULT 10
UBIG_PASMU
ID UBIG_PASMU STANDARD; PRT; 242 AA.
AC 09CM16;

DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 3-demethylubiquinone-9-3-methyltransferase (EC 2.1.1.64) (3,4-
DE dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHNB
DE methyltransferase).
GN UBIG OR PM0840.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang O., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- CARBOLYTIC ACTIVITY: S-adenosyl-L-methionine + 3-
CC -1- demethylubiquinone-9 = S-adenosyl-L-homocysteine + ubiquinone-9.
CC -1- PATHWAY: Ubiquinone biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE UBIG/COQ3 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AE006122; AK02924.1;
CC InterPro: IPR001601; Methyltransf.
CC InterPro: IPR000051; SAM_bind.
CC Ubiquinone biosynthesis; Transferase; Methyltransferase;
CC Complete proteome.
CC KW
CC SO SEQUENCE 242 AA; 27284 MW; B6318FA1ACC01BDD CRC64;

Query Match 7.3%; Score 134.5; DB 1; Length 242;
Best Local Similarity 24.9%; Pred. NO. 0.00038;
Matches 52; Conservative 37; Mismatches 75; Indels 45; Gaps 11;
OY 116 LPFASVSEERSKMP-KSIYDVGGCIGSSRYLAKKFGATSVGITLS--PYOAOANALAA 172
DB 35 LRLSYIAAQANGLTGKRVLDVGGCGGLSESMKQ--GAYVTGIDMSAPLQVARKALPS 93
OY 173 AGCL-ADKVSFOVADALQOPFS-----DGOVDLWMSHESGSHMDAKFYGELARVAA 224
DB 94 --GLHIDYQOITTEFLQNOTALFAERGEDKFDVITCMELEHVPDPSSIAACKQQLK 151
OY 225 PGATIIIVTWCHRDLPDEQSLHPV-----EODLLAKKI-----CDAYYLPV---WCS 268
DB 152 PNVGIVFST-----INRTLKAMALVIIGAYVYKMLPKRGTHDYDKFKIRPELLHMD 203
OY 269 TSDYVKLLQSLSLQDKSEDMRSFVAPFW 297
DB 204 EA-----QLTCLDMVGYHYNPLTGKFW 225

RESULT 11
UBIG_NEIMB
ID UBIG_NEIMB STANDARD; PRT; 238 AA.
AC 09JX17;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 3-demethylubiquinone-9-3-methyltransferase (EC 2.1.1.64) (3,4-
DE dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHNB
DE methyltransferase).
GN UBIG OR NMB2030.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;

```

[1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=MC58 / Serogroup B;
 RC MEDLINE=20175755; PubMed-10710307;
 RA Tettein H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Elsen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 RA Nelson W.C., Gwin M.L., DeBoy R., Peterson J.D., Hickey E.K.,
 RA Hafiz D., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Clecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
 RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
 RA Gill J., Scarlato V., Maignani V., Piza M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappelli R., Venter J.C.;
 RT *Complete genome sequence of *Neisseria meningitidis* serogroup B strain
 MC58.
 RL Science 287:1809-1815(2000).
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-
 CC demethylubiquinone-9 -> S-adenosyl-L-homocysteine + ubiquinone-9.
 CC -1- SIMILARITY: Ubiquinone biosynthesis.
 CC -1- SIMILARITY: BELONGS TO THE UBIG/COQ3 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC DR EMBL: AE002552; AAF42352.1; ALT_INIT.
 DR TIGR: NMB2030;
 DR InterPro: IPR001601; Methyltransf.
 DR InterPro: IPR000051; SAM_bind.
 KW Ubiquinone biosynthesis; Transferase; Methyltransferase;
 KM Complete proteome.
 SQ SEQUENCE 238 AA; 26529 MW; 7CA7305A824BADFE CRC64;

Query Match 7.2%; Score 131.5; DB 1; Length 238;
 Best Local Similarity 27.1%; Pred. No. 0.00066;
 Matches 46; Conservative 25; Mismatches 64; Indels 35; Gaps 6;

QY 130 KSIIVVCGIGIGSSRYLAKFGATSVGTLSPVOAORANALAAOGLADKVSFO---VAD 186
 DB 53 KAVLDVCGCGGILAESMARGAFAVKGIDMAQOSLETALHAAALNNVAD-IEECIRVAD 111
 QY 187 -ALQPFSDGDFDLVSMESGEMDPKAFVGEELARVAAPGAIITVTWCHRDLPDEOS 245
 DB 112 LAEAPHS---EDVYTCMEHMEHVDPALIVACANLVKPGDWFEPT----- 156
 QY 246 LHPWODLLKRICDAVYLPKWCSTSDYVKLLQSLQDIKSEDMSEFVAP 295
 DB 157 INKNPKSYLHLIVAAEYL-----LKFVPRKGTIDMKKFIAP 191

RESULT 12

UBIG_RICCN
 ID UBIG_RICCN STANDARD: PRT; 289 AA.
 AC Q92807;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 3-demethylubiquinone-9-3-methyltransferase (EC 2.1.1.64) (3,4-
 DE dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHMB
 DE methyltransferase).
 GN UBIG OR RC0965.
 OS Rickettsia conorii.
 CC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 CC Rickettsiaceae; Rickettsiasee; Rickettsia.
 OX NCBI_TaxID=781;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Malish 7;
 RX MEDLINE=21442074; PubMed-11557893;
 RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,

RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
 RA Raoult D.;
 RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii";
 RL Science 293:2093-2098(2001).
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-
 CC demethylubiquinone-9 -> S-adenosyl-L-homocysteine + ubiquinone-9.
 CC -1- SIMILARITY: Ubiquinone biosynthesis.
 CC -1- SIMILARITY: BELONGS TO THE UBIG/COQ3 FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 REPEAT INSERT DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboratio
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on it
 CC use by non-profit institutions as long as its content is in no wa
 CC modified and this statement is not removed. Usage by and for commercia
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC DR EMBL: AE008649; AAL03503.1;
 DR InterPro: IPR001601; Methyltransf.
 DR InterPro: IPR000051; SAM_bind.
 DR TIGR: TIGR01045; RPE: 1.
 KW Ubiquinone biosynthesis; Transferase; Methyltransferase;
 KM Complete proteome.
 FT DOMAIN 50 REPEAT INSERT.
 SQ SEQUENCE 289 AA; 33074 MW; 4CAEA6827EC7B977 CRC64;

Query Match 7.2%; Score 131.5; DB 1; Length 289;
 Best Local Similarity 22.5%; Pred. No. 0.00085;
 Matches 60; Conservative 51; Mismatches 91; Indels 65; Gaps 12

QY 61 DKKKLQKGAIEFYDESSGLMENIWGDH-----MHGCFYDSST--- 98
 DB 5 DKKELEK-----FEKISHMMWKKDGEFGLHRINPRLLEYIEKITTHVRHLSKLYRE 59
 QY 99 -VSLSDHRAQIRMIQE-----SLRFAVSERSRKSPSYDVCGIGG--S 142
 DB 60 ELVGMNQHSTAYALVAREPASSRLTKRLPLEAEFEKMSNDISK--LEILDVCGG-GGLIA 116
 QY 143 SHYLAKKFGATSVGTLSPVOAORANALAAOGLADKVSFOVDALQPFSDGDFLWMS 202
 DB 117 TPLAAGFVNTALDQSNIEF--ATAVAKENY--KINK-LOSTIEELDSKLYDVIC 171
 QY 203 MESGEHMPDKAFVGEELARVAAPGAIITVTWCHRDLPDEOSLHPEWODLLKRICDAY 262
 DB 172 LEVIEHVENVOQFILLVHKIRPNGMAIIST-----INTRKAYILGIIVAEY 219
 QY 263 LPKWC--STSDYVKLLQSLQDIKSE 287
 DB 220 ILGWVPKNTHDYSKFLKPLEIYEMLD 246

RESULT 13

YE03_MYCTU
 ID YE03_MYCTU STANDARD: PRT; 274 AA.
 AC P71671;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein RVI403c precursor.
 GN RVI403C OR MT1447 OR MYCY21B4.20C.
 OS Mycobacterium tuberculosis.
 CC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 CC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed-96344230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. II, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,

```

RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sultun J.E., Taylor K., Whitehead S., Barrell B.G.,
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CCDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolony J.F., Nelson W.C., Umayan L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weldman J., Khouri H., Gill J., Mikula A.,
RA Bhalal W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: STRONG, TO M.TUBERCULOSIS RV1405C AND SOME, TO
CC R.SPHEROIDES PMTA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; Z80108; CAB02182.1;
DR EMBL; AE007015; AAK45712.1;
DR TIGR; MT1447;
DR Tuberculist; RV1403c;
DR InterPro: IPR000051; SAM bind.
KW Hypothetical protein; Signal; Complete proteome.
FT SIGNAL 1 30
FT CHAIN 31 274
FT SEQUENCE 274 AA; 30077 MW; B554838D60BDC0 CRC64;
SQ
Query Match 7.1%; Score 130; DB 1; Length 274;
Best Local Similarity 30.3%; Pred. No. 0.0011;
Matches 53; Conservative 19; Mismatches 79; Indels 24; Gaps 6;
QY 132 IVDVCGGIGSSRYLAKKFGATSVGITLSPVQAQRANALAAAGLADKVSFQVADALQOP 191
DB 53 VLDVNAAGSGNVS-IPAAAGAHVHTASDLTPPELLRAQAARAAAGL-ELGMEANAELALP 109
QY 192 FSDGQFDLVWMESEGHMPDKAKFVGEIARVAAPCAIIITVWCHRD-LGPDOSLHPWE 250
DB 110 FSAGFEDAVLSTIGVMAFRHQRADLAVRCRGGKISTLWMPDEGFYGLKSLTIRPYR 169
QY 251 QDLKKICDAVYLPVPA-----WCSTSDYVKLLQSLQSLQDIKSEDSM-----RFVAP 295
DB 170 PT-----LPAAGFHEVWMSGSEDIYVSGLFKRDVSDIRKRGSLTVDRFGCP 214
RESULT 14
UBIG_NEIMA STANDARD; PRT; 238 AA.
AC Q9JME6;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 3-demethylubiquinone-9 3-methyltransferase (EC 2.1.1.64) (3,4-
DE dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHMB
DE methyltransferase).
CN UBIG OR NAA0410.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCB1_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22491 / Serogroup A / Serotype 4A;

```

```

RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis 22491."
RL Nature 404:502-506(2000).
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3-
CC 3-demethylubiquinone-9 - S-adenosyl-L-homocysteine + ubiquinone-9.
CC -1- PATHWAY: Ubiquinone biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE UBIG/COO3 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AL162753; CAB83709.1;
DR InterPro: IPR001601; Methyltransf.
DR InterPro: IPR000051; SAM bind.
KW Ubiquinone biosynthesis; Transferase; Methyltransferase;
KW Complete proteome.
SQ SEQUENCE 238 AA; 26571 MW; 629D6F6ED3BA8AA CRC64;
Query Match 7.1%; Score 129.5; DB 1; Length 238;
Best Local Similarity 26.5%; Pred. No. 0.00097;
Matches 45; Conservative 26; Mismatches 64; Indels 35; Gaps 6;
QY 130 KSIYDVCGGIGSSRYLAKKFGATSVGITLSPVQAQRANALAAAGLADKVSFQVADALQOP 186
DB 53 KRVLDVCGGIGGGLAESMARCAAFYKIDMAEOSLETARLHAALNVAD-TEYECIRVED 111
QY 187 -ALQOPFSDGQFDLVWMESEGHMPDKAKFVGEIARVAAPCAIIITVWCHRD-LGPDQS 245
DB 112 LAEAPHS---FDVYTCKEMHEHPDPAPAIYRACKLVKPGMGFFST----- 156
QY 246 LHPWQDLKKICDAVYLPVPAWCSTSDYVKLLQSLQSLQDIKSEDSMRFVAP 295
DB 157 INKNKSTLHLIVAAYL-----LKFVKGTHDMKFFIVP 191
RESULT 15
UBIG_PSEAE STANDARD; PRT; 232 AA.
AC Q9H263;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 3-demethylubiquinone-9 3-methyltransferase (EC 2.1.1.64) (3,4-
DE dihydroxy-5-hexaprenylbenzoate methyltransferase) (DHMB
DE methyltransferase).
CN UBIG OR PA3171.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCB1_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warriner P.,
RA Hickey M.J., Goltzman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Gager R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an

```

```

RT opportunistic pathogen."
RL Nature 406:959-964(2000).
CC -1- CATALYTIC ACTIVITY: S'-adenosyl-L-methionine + 3-
CC demethylubiquinone-9 = S'-adenosyl-L-homocysteine + ubiquinone-9.
CC -1- PATHWAY: Ubiquinone biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE UBIQ/COQ3 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE004741; AAC06559.1;
CC InterPro: IPR001601; Methyltransf.
CC InterPro: IPR000051; SAM_bind.
CC InterPro: IPR004033; UbiE/COQ5_Metrf.
CC Pfam; PF01209; UbiE_methyltran; 1.
CC Ubiquinone biosynthesis; Transferase; Methyltransferase;
CC Complete proteome.
CC SEQUENCE 232 AA; 25859 MW; 643EA81D90A4177C CRC64;
SQ

```

Query Match 7.0%; Score 128; DB 1; Length 232;
 Best Local Similarity 29.1%; Pred. No. 0.0013;
 Matches 43; Conservative 27; Mismatches 60; Indels 18; Gaps 7;

```

OY 90 HGEYDSSTVS-LSDHRAQRIEMIOESLRPASVSEERSKWPMSIYDVCGGIGGSSRYLAK 148
   |::| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 17 HRWMDRESEFKPLNDINPLRYVWIDERAGLAG-----KKVIDIGGGGILSEAMNQ 67

```

OY 149 KEGATVCGITL--SPVQAQRANALAAOGLADKVSFOVADALQOPFS--DGOFDLVSMES 205
 :||: ||: :|: | | | | | | | | | | | | | | | | | | | | | | | |
Db 68 R-GASVTCIDMGEPALVAVRLHQLES--GVA--VDYROIATFQMAEMEPGQFDVYTCLEM 122

OY 206 GEHMPDKAKFVGEIARVAPGAIITITVT 233
 ||: || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 123 LEHVDPDPASVIRACHRLVKGGOVPLST 150

Search completed: March 18, 2003, 17:38:08
 Job time : 14 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	1159	63.3	369	10	Q8VWX2	Q8VWX2 perilla fru
2	1154.5	63.1	348	10	Q9XIP9	Q9XIP9 arabidopsis
3	1151.5	62.9	348	10	Q9SZK1	Q9SZK1 arabidopsis
4	659	36.0	280	16	Q8WY13	Q8WY13 anabaena sp
5	647.5	35.4	317	16	Q55809	Q55809 synecocyst
6	392.5	21.4	330	16	Q8VY60	Q8VY60 anabaena sp
7	392	21.4	318	16	P74388	P74388 synecocyst
8	319	17.4	416	3	Q9C2D6	Q9C2D6 neutrospora
9	280.5	15.3	283	2	Q9X5C9	Q9X5C9 streptomyce
10	277	15.1	272	2	Q52570	Q52570 amycolatops
11	261.5	14.3	278	2	Q9EY12	Q9EY12 streptomyce
12	247	13.5	352	10	Q24153	Q24153 nicotiana t
13	241	13.2	363	10	Q41586	Q41586 triticum ae
14	239	13.1	363	10	Q41587	Q41587 triticum ae
15	238	13.0	283	2	Q9S0N6	Q9S0N6 streptomyce
16	234.5	12.8	264	16	Q96FP8	Q96FP8 rhizobium l

17	233.5	12.7	565	2	0.9KJ10	0.9K120	actinopolys
18	227	12.4	344	10	P93852	zeae may's (m	
19	226	12.3	279	2	0.9KJ21	0.9K121	ectothiorino
20	226	12.3	349	10	0.94US4	0.9Jf54	arabidopsis
21	225.5	12.3	359	10	0.82426	0.82426	oryza sativ
22	225	12.3	344	10	0.94215	0.94215	zea may's (m
23	223	12.2	359	10	0.23655	0.23655	arabidopsis
24	221.5	12.1	357	10	0.24154	0.24154	nicotiana t
25	220.5	12.0	275	2	0.9ALM7	0.9ALM7	saccharopol
26	213	11.6	280	16	0.98K44	0.98K44	rhizobium l
27	212.5	11.6	361	10	0.39227	0.39227	arabidopsis
28	212.5	11.6	363	10	0.9LW25	0.9LW25	arabidopsis
29	210.5	11.5	361	10	0.82427	0.82427	oryza sativ
30	210	11.5	379	3	0.9P3R1	0.9P3R1	neurospora
31	209	11.4	377	3	0.96M42	0.96M42	pneumocyst
32	206.5	11.3	346	10	0.82720	0.82720	nicotiana t
33	204	11.1	336	10	0.43445	0.43445	glycine max
34	203.5	11.1	367	10	0.94M02	0.94M02	arabidopsis
35	202.5	11.1	275	2	0.9X508	0.9X508	streptomyce
36	199	10.9	359	3	0.9C439	0.9C439	pneumocyst
37	194.5	10.6	346	10	0.24C28	0.24C28	ricinus com
38	192.5	10.5	349	10	0.82434	0.82434	nicotiana t
39	190	10.4	374	10	0.9LPE3	0.9LPE3	arabidopsis
40	190	10.4	491	10	0.9FR44	0.9FR44	arabidopsis
41	190	10.4	295	10	0.9LW13	0.9LW13	arabidopsis
42	188	10.3	498	10	0.9C9V1	0.9C9V1	arabidopsis
43	188	10.3	555	10	0.9C6B9	0.9C6B9	arabidopsis
44	187	10.2	475	10	0.944H0	0.944H0	arabidopsis
45	184	10.1	491	10	0.94XH3	0.94XH3	lycopersico

ALIGNMENTS

RESULT 1

ID Q8VWX2 PRELIMINARY; PRT; 369 AA.

DT 01-MAR-2002 (TREMblrel_20, Created)
DT 01-MAR-2002 (TREMblrel_20, Last sequence update)
DT 01-JUN-2002 (TREMblrel_21, Last annotation update)
DE Gamma-tocopherol methyltransferase.
GN TMF.
OS *Perilla frutescens*.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Lamiales; Lamiaceae; Perilla.
OX NCBI_TaxID=48386;
RX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=CV, OKDONG; TISSUE=DEVELOPING SEED;
RC Kim K.-H., Hwang S.-K., Hwang Y.-S.;
RA *Cloning of Perilla gamma-tocopherol methyltransferase*;
RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF213481; FAL36993.1; -
DR InterPro: IPR000051; SAM_1bind.
KW Methyltransferase; Methyltransferase.
SW SEQUENCE 369 AA; 40739 MW; 2748E77219617E9D CRC64;

[illegible]

[illegible]

OY	32	GPBSWAPIRASAASE-----RCGTVLEQKPKKDKKKLCKGIAEFYDESSGLMWTNG	85
Db	23	GSKSLFRSSSSSSSVSMTTTRGNVAVAATAATST--ALRKGIAEFNETSGLMEEITWG	80
OY	86	DHMHGFGYDSSTVSLSD--HRAAQIRMIQESLRFASVS-EERSKMPKSIYDVCGGIGS	142
Db	81	DHMHGFGYDPPSSVGLSDSGHKEAQIRMI EESLRAGVTDDEEBEKKIKKVVYDVGCGIGS	140
OY	143	SRYLAKKRGATSVGTTLSPVQAQRANALAAAGLADKYSFOYADALDQPFSDGQFDLWVS	202
Db	141	SRYLASKFGAGCIGITTLSPVQAKRANDLAAAGSLAHKSKASFOYADALDQPFEDGKFDLWVS	200
OY	203	MESGHHMPDKKAFVGEELRVAPAGALIIIVTWCCHRDLDGPEQSLPMEQDILKRTCDAYY	262
Db	201	MESGHHMPDKKAFVVELRVAPAPGGRITIIIVTWCCHNLNSGEBALQPMQONILDKICKIFY	260
OY	263	LPAMCSTDYVKLLQSLSLQDIKSDMSRFVAFPMFPAVIRSAFTWKGSLSSLSGQKTIK	322
Db	261	LPAMCSTDYVNLQSLSLQDIKCADMSEBNVAFPMFPAVIRALTWKGSLVLLRSQMSIK	320
OY	323	GALAMPIMEGKKDKLIRFAITTCCKP 349	
Db	321	GALTPMLIMEGKKGVIRKFGIITCCKP 347	
RESULT 3			
O9ZSKI	ID	PRELIMINARY; PRT; 348 AA.	
AC	O9ZSKI		
DT	01-MAY-1999 (TREMBUREL 10, Created)		
DT	01-MAY-1999 (TREMBUREL 10, Last sequence update)		
DT	01-MAR-2002 (TREMBUREL 20, Last annotation update)		
DE	Gamma-tocopherol methyltransferase.		
GN	G-TMT.		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota, Viridiplantae: Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots; Rosidae;		
CC	eucosids II; Brassicales; Brassicaceae; Arabidopsis.		
OX	NCBI_Taxid:3702;		
RN	(1)		
RP	SEQUENCE FROM N.A.		
RA	Shintani D.K., Dellapenna D.;		
RT	"Elevating Vitamin E content of plants through metabolic		
RT	engineering.";		
RL	Science 0:0-0(1998)		
DR	EMBL; AF104220; AAD02882.1; -		
DR	InterPro: IPR001601; Methyltransf.		
DR	InterPro: IPR000051; SAM bind.		
KW	Methyltransferase; Transferase.		
SO	SEQUENCE 348 AA; 38091 MW; 44DB18A722E0725F CnC64;		
Query Match 62.9%; Score 1151.5; DB 10; Length 348;			
Best Local Similarity 69.1%; Pred. No. 1.8e-86;			
Matches 226; Conservative 29; Mismatches 61; Indels 11; Gaps 4;			
OY	32	GPBSWAPIRASAASE-----RCGTVLEQKPKKDKKKLCKGIAEFYDESSGLMWTNG	85
Db	23	GSKSLFRSSSSSSSVSMTTTRGNVAVAATAATST--ALRKGIAEFNETSGLMEEITWG	80
OY	86	DHMHGFGYDSSTVSLSD--HRAAQIRMIQESLRFASVS-EERSKMPKSIYDVCGGIGS	142
Db	81	DHMHGFGYDPPSSVGLSDSGHKEAQIRMI EESLRAGVTDDEEBEKKIKKVVYDVGCGIGS	140
OY	143	SRYLAKKRGATSVGTTLSPVQAQRANALAAAGLADKYSFOYADALDQPFSDGQFDLWVS	202
Db	141	SRYLASKFGAGCIGITTLSPVQAKRANDLAAAGSLAHKSKASFOYADALDQPFEDGKFDLWVS	200
OY	203	MESGHHMPDKKAFVGEELRVAPAGALIIIVTWCCHRDLDGPEQSLPMEQDILKRTCDAYY	262
Db	201	MESGHHMPDKKAFVVELRVAPAPGGRITIIIVTWCCHNLNSGEBALQPMQONILDKICKIFY	260
OY	263	LPAMCSTDYVKLLQSLSLQDIKSDMSRFVAFPMFPAVIRSAFTWKGSLSSLSGQKTIK	322
Db	261	LPAMCSTDYVNLQSLSLQDIKCADMSEBNVAFPMFPAVIRALTWKGSLVLLRSQMSIK	320
OY	323	GALAMPIMEGKKDKLIRFAITTCCKP 349	
Db	321	GALTPMLIMEGKKGVIRKFGIITCCKP 347	
OY	263	LPAMCSTDYVKLLQSLSLQDIKSDMSRFVAFPMFPAVIRSAFTWKGSLSSLSGQKTIK	322
Db	261	LPAMCSTDYVNLQSLSLQDIKCADMSEBNVAFPMFPAVIRALTWKGSLVLLRSQMSIK	320
OY	203	MESGHHMPDKKAFVGEELRVAPAGALIIIVTWCCHRDLDGPEQSLPMEQDILKRTCDAYY	262
Db	201	MESGHHMPDKKAFVVELRVAPAPGGRITIIIVTWCCHNLNSGEBALQPMQONILDKICKIFY	260
OY	143	SRYLAKKRGATSVGTTLSPVQAQRANALAAAGLADKYSFOYADALDQPFSDGQFDLWVS	202
Db	141	SRYLASKFGAGCIGITTLSPVQAKRANDLAAAGSLAHKSKASFOYADALDQPFEDGKFDLWVS	200
OY	203	MESGHHMPDKKAFVGEELRVAPAGALIIIVTWCCHRDLDGPEQSLPMEQDILKRTCDAYY	262
Db	201	MESGHHMPDKKAFVVELRVAPAPGGRITIIIVTWCCHNLNSGEBALQPMQONILDKICKIFY	260
OY	263	LPAMCSTDYVKLLQSLSLQDIKSDMSRFVAFPMFPAVIRSAFTWKGSLSSLSGQKTIK	322
Db	261	LPAMCSTDYVNLQSLSLQDIKCADMSEBNVAFPMFPAVIRALTWKGSLVLLRSQMSIK	320
OY	203	MESGHHMPDKKAFVGEELRVAPAGALIIIVTWCCHRDLDGPEQSLPMEQDILKRTCDAYY	262
Db	201	MESGHHMPDKKAFVVELRVAPAPGGRITIIIVTWCCHNLNSGEBALQPMQONILDKICKIFY	260
OY	143	SRYLAKKRGATSVGTTLSPVQAQRANALAAAGLADKYSFOYADALDQPFSDGQFDLWVS	202
Db	141	SRYLASKFGAGCIGITTLSPVQAKRANDLAAAGSLAHKSKASFOYADALDQPFEDGKFDLWVS	200
OY	203	MESGHHMPDKKAFVGEELRVAPAGALIIIVTWCCHRDLDGPEQSLPMEQDILKRTCDAYY	262
Db	201	MESGHHMPDKKAFVVELRVAPAPGGRITIIIVTWCCHNLNSGEBALQPMQONILDKICKIFY	260
OY	263	LPAMCSTDYVKLLQSLSLQDIKSDMSRFVAFPMFPAVIRSAFTWKGSLSSLSGQKTIK	322
Db	261	LPAMCSTDYVNLQSLSLQDIKCADMSEBNVAFPMFPAVIRALTWKGSLVLLRSQMSIK	320
OY	203	MESGHHMPDKKAFVGEELRVAPAGALIIIVTWCCHRDLDGPEQSLPMEQDILKRTCDAYY	262
Db	201	MESGHHMPDKKAFVVELRVAPAPGGRITIIIVTWCCHNLNSGEBALQPMQONILDKICKIFY	260
OY	143	SRYLAKKRGATSVGTTLSPVQAQRANALAAAGLADKYSFOYADALDQPFSDGQFDLWVS	202
Db	141	SRYLASKFGAGCIGITTLSPVQAKRANDLAAAGSLAHKSKASFOYADALDQPFEDGKFDLWVS	200
OY	203	MESGHHMPDKKAFVGEELRVAPAGALIIIVTWCCHRDLDGPEQSLPMEQDILKRTCDAYY	262

DB 261 LKMGSTDDYVNLDSHSLQDICKADMSVNAFPAVIRLTALTKGLVSLRSKSTIK 320

OY 323 GALAMPLEMGYKKDLKFAITTCRNP 349

DB 321 GALTPLMIEGKGVKGVITTCORP 347

RESULT 4

OBYW13 PRELIMINARY: PRT: 280 AA.

AC 08YW13: PRELIMINARY: PRT: 280 AA.

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Gamma-tocopherol methyltransferase.

GN ALR1803.

OS Anabaena sp. (strain PCC 7120).

OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.

OX NCBI_TaxID=103690.

RN (1)

RP SEQUENCE FROM N.A.

RX MEDLINE=21595285; PubMed=11759840;

RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,

RA Matanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,

RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,

RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,

RA Yasuda M., Tabata S.;

RT *Complete genomic sequence of the filamentous nitrogen-fixing

RT cyanobacterium Anabaena sp. strain PCC 7120.;

RL DNA Res. 8:205-213(2001).

DR EMBL: AP003587; BAB73502.1;

DR InterPro: IPR001601; Methyltransf.

DR InterPro: IPR000051; SAM_bind.

DR Pfam: PF01209; UblE_methyltran; 1.

KW Transferase: Methyltransferase; Complete proteome.

SO SEQUENCE 280 AA; 31309 MW; 663FD841294B9133 CRC64;

Query Match 36.0%; Score 659; DB 16; Length 280;

Best local similarity 49.3%; Pred. No. 3.9e-46;

Matches 138; Conservative 42; Mismatches 92; Indels 8; Gaps 3;

OY 65 LKMGSTDDYVNLDSHSLQDICKADMSVNAFPAVIRLTALTKGLVSLRSKSTIK 124

DB 5 LKMGSTDDYVNLDSHSLQDICKADMSVNAFPAVIRLTALTKGLVSLRSKSTIK 60

OY 125 RSKWPKSYDVGGCGIGSSRYLAKKFGATSVGTLSPVQARANALAAOGLADKVSFOV 184

DB 61 --QTAENILDVGGCGIGSSRYLAKKFGATSVGTLSPVQARANALAAOGLADKVSFOV 118

OY 185 ADALQOPFSDGQFDLVWMSGEMHPDKAKFVGLARVAAGALIIITVCHRDLPDEQ 244

DB 119 ANQAMPFDNSFDLVWMSGEMHPDKAKFVGLARVAAGALIIITVCHRDLPDEQ 176

OY 245 SLHPWEDDLKKICDAIYLPAMCSTSDYVNLDSHSLQDICKADMSVNAFPAVIRSA 304

DB 177 PLRADEKRLIEDIYRVCLEPVISLPEYEAIRQLPLNNRTADMSVNAFPAVIRSA 236

OY 305 FTWKGLSSLSGCKTKGALAMPLEMGYKKDLKFAITTCRNP 344

DB 237 FTWQALFGLRAGWTITGALSLGALMRGTERGLIRFGL 276

RESULT 5

O55809 PRELIMINARY: PRT: 317 AA.

AC 055809: PRELIMINARY: PRT: 317 AA.

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

DE DELTA(24)-sterol C-methyltransferase.

GN ERG6 OR SLR0089.

OS Synechocystis sp. (strain PCC 6803).

OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.

OX NCBI_TaxID=1148;

RN (1)

RP SEQUENCE FROM N.A.

RX MEDLINE=96127529; PubMed=8590279;

RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,

RA Sugiyama M., Tabata S.;

RT *Sequence analysis of the genome of the unicellular cyanobacterium

RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb

RT region from map positions 64k to 92k of the genome.;

RL DNA Res. 2:153-166(1995).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=97061201; PubMed=8905231;

RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,

RA Miyajima N., Hiroseawa M., Sugiyama M., Sasamoto S., Kimura T.,

RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,

RA Shimo S., Takeuchi C., Wada T., Matanabe A., Yamada M., Yasuda M.,

RA Tabata S.;

RT *Sequence analysis of the genome of the unicellular cyanobacterium

RT Synechocystis sp. strain PCC6803. II. Sequence determination of the

RT entire genome and assignment of potential protein-coding regions.;

RL DNA Res. 3:109-136(1996).

DR EMBL: D64004; BAA10562.1;

DR InterPro: IPR001601; Methyltransf.

DR InterPro: IPR000051; SAM_bind.

KW Transferase: Methyltransferase; Complete proteome.

SO SEQUENCE 317 AA; 35571 MW; 98DD01C081B89F72 CRC64;

Query Match 35.4%; Score 647.5; DB 16; Length 317;

Best local similarity 47.4%; Pred. No. 4.1e-45;

Matches 135; Conservative 45; Mismatches 102; Indels 3; Gaps 3;

OY 65 LKMGSTDDYVNLDSHSLQDICKADMSVNAFPAVIRLTALTKGLVSLRSKSTIK 124

DB 34 LYKRIKFFYDSSGLMEDVGMHNGHYGPHGYRI-DRROAIDILKELAWA-VPO 91

OY 125 RSKWPKSYDVGGCGIGSSRYLAKKFGATSVGTLSPVQARANALAAOGLADKVSFOV 184

DB 92 SAR-PKRIIDLGCGIGSSRYLAKKFGATSVGTLSPVQARANALAAOGLADKVSFOV 150

OY 185 ADALQOPFSDGQFDLVWMSGEMHPDKAKFVGLARVAAGALIIITVCHRDLPDEQ 244

DB 151 ANALDLFPASDFPWSLSGEMHPDKAKFVGLARVAAGALIIITVCHRDLPDEQ 210

OY 245 SLHPWEDDLKKICDAIYLPAMCSTSDYVNLDSHSLQDICKADMSVNAFPAVIRSA 304

DB 211 PLRADEKRLIEDIYRVCLEPVISLPEYEAIRQLPLNNRTADMSVNAFPAVIRSA 270

OY 305 FTWKGLSSLSGCKTKGALAMPLEMGYKKDLKFAITTCRNP 349

DB 271 FDRVILALQAGRIINALCLRLMKWYRGVGLGIRFGL 315

RESULT 6

O8YV60 PRELIMINARY: PRT: 330 AA.

AC 08YV60: PRELIMINARY: PRT: 330 AA.

DT 01-MAR-2002 (TREMBlrel. 20, Created)

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Hypothetical protein A112121.

GN A112121.

OS Anabaena sp. (strain PCC 7120).

OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.

OX NCBI_TaxID=103690;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21595285; PubMed=11759840;

RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,

RA Matanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,

RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,

